The slide features a white background with abstract, colorful geometric shapes in the corners. On the left, there are overlapping triangles in shades of green, blue, orange, and purple. On the right, there are larger, more complex shapes in green, blue, purple, and red. The title 'Recommendation Engine' is centered in a brown, sans-serif font.

# Recommendation Engine

Heather Ciallella, Selina Hui, Sachin Saxena

# Our Task

- Proof of Concept for an Event Recommendation Engine
- Start with Zipped Raw Data
- End with at Least 5 Recommendations on a Webpage

TICKETLEAP

How it works

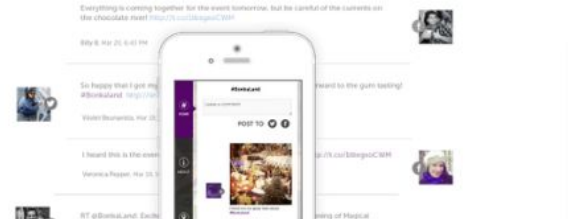
Pricing

Customers

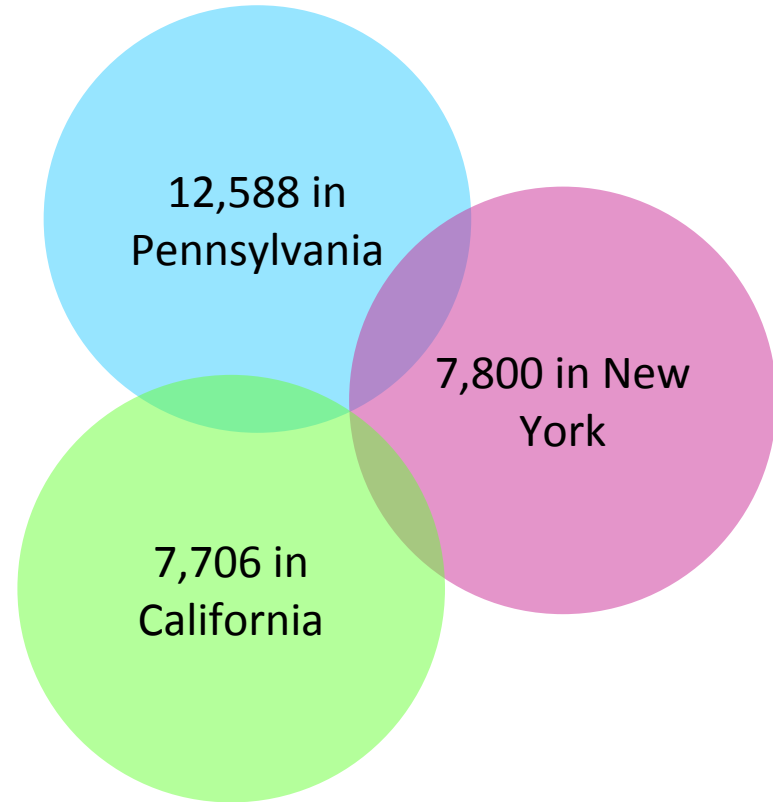
Events U

Help Center

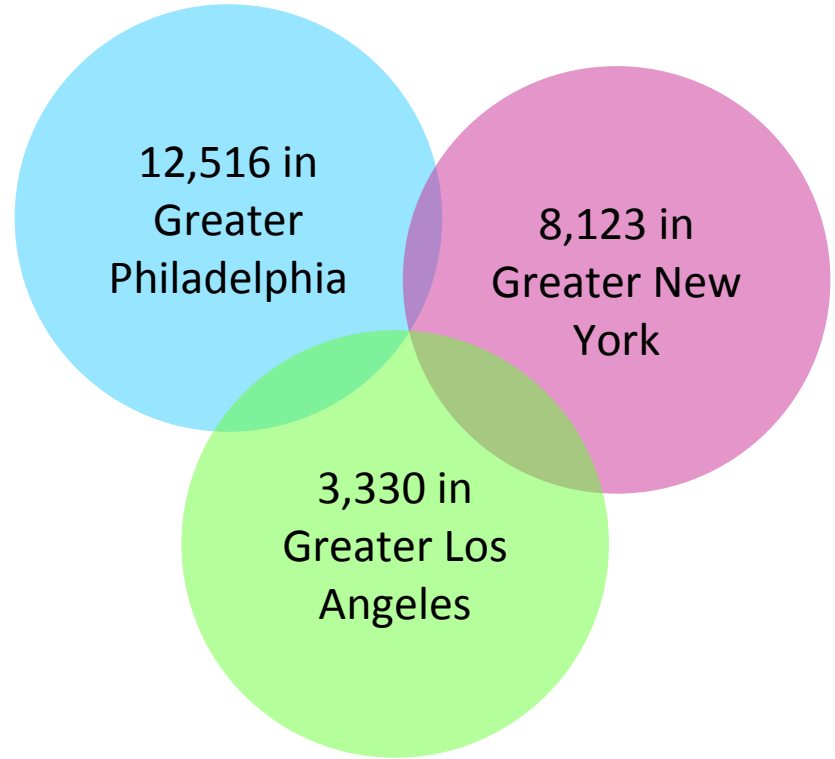
Sell tickets online to your ever



## Our Data

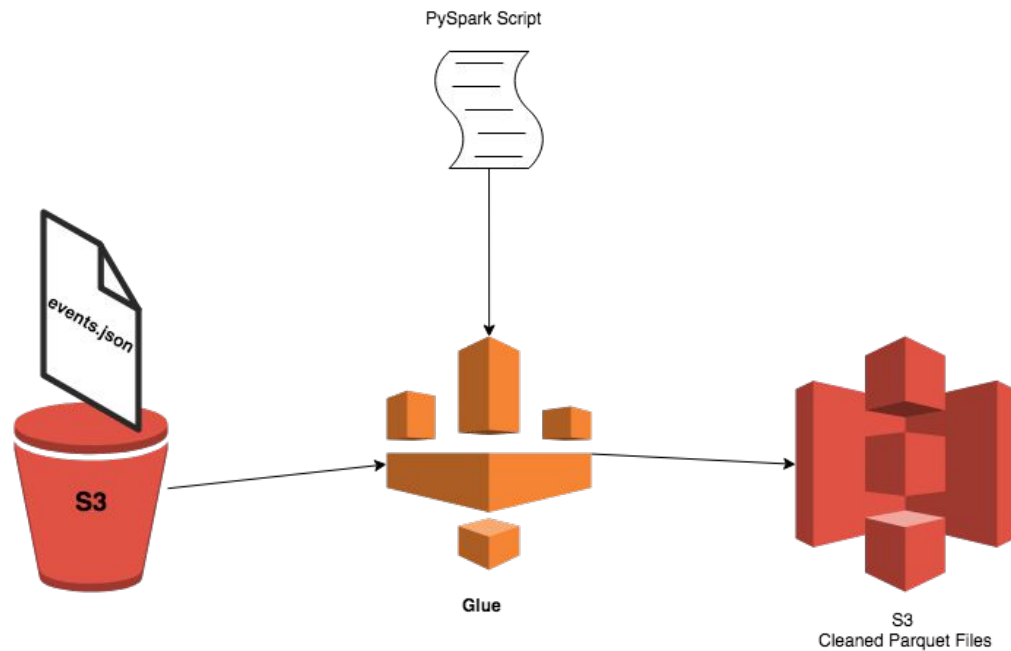


## Our Data



# Phase 1

## Pipeline



# Clean and Normalize

```
event_description    Night of the senses;Ã an exclusive gala benefitingÃ the ArtesanoGroup
Foundation.\n\nAn engaging Dinner and Silent Auction inspired by the cookbook âLos
Alimentos del Deseoâ by the Venezuelan Chef and JournalistÃ Maruja Dagnino.\n\nJoin
Special guests Boris Izaguirre, George Harris, and Dr. Nancy Alvarez as they lead you
through this journey with author Maruja Dagnino.\n\nMiami Beach Botanical GardenÃ 2000
Convention CenterÃ \n\n07:30pm to 11:00pm Red Carpet 07:00pm\n\nÃ \n
event_end_utc        10 Feb 2018 04:00:00
event_name           A night of the senses
event_start_utc       10 Feb 2018 00:00:00
facebook_event_id
hashtag              None
organization_id       la-fontana-ristorante
organization_name     La Fontana Ristorante
tags                  []
venue_city            Miami Beach
venue_name            Miami Beach Botanical Garden
venue_state           FL
venue_street          2000 Convention Center Dr
venue_timezone        US/Eastern
venue_zip             33139
Name: cdd7d714-f260-11e7-9897-22000a9b28c5
```

# Clean and Normalize

Night of the senses;Â an exclusive gala benefitingÂ the ArtesanoGroup Foundation.

An engaging Dinner and Silent Auction inspired by the cookbook âLos Alimentos del Deseoâ by the Venezuelan Chef and JournalistÂ Maruja Dagnino.

Join Special guests Boris Izaguirre, George Harris, and Dr. Nancy Alvarez as they lead you through this journey with author Maruja Dagnino.

Miami Beach Botanical GardenÂ 2000 Convention CenterÂ

07:30pm to 11:00pm Red Carpet 07:00pm

Â

Night of the senses; an exclusive gala benefiting the ArtesanoGroup Foundation. An engaging Dinner and Silent Auction inspired by the cookbook Los Alimentos del Deseo by the Venezuelan Chef and Journalist Maruja Dagnino. Join Special guests Boris Izaguirre, George Harris, and Dr. Nancy Alvarez as they lead you through this journey with author Maruja Dagnino. Miami Beach Botanical Garden 2000 Convention Center 07:30pm to 11:00pm Red Carpet 07:00pm

# Cleaned

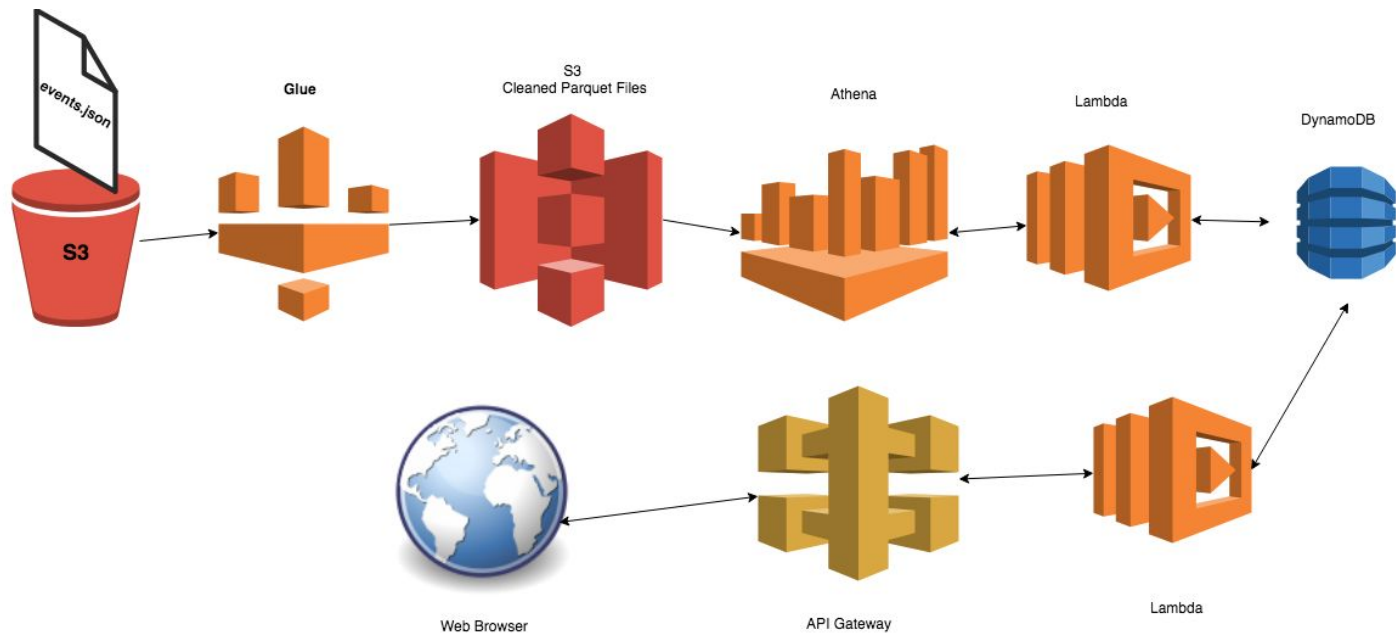
```
event_description Night of the senses;Â an exclusive gala benefitingÂ the ArtesanoGroup
Foundation.\n\nAn engaging Dinner and Silent Auction inspired by the cookbook âLos
Alimentos del Deseoâ by the Venezuelan Chef and JournalistÂ Maruja Dagnino.\n\nJoin
Special guests Boris Izaguirre, George Harris, and Dr. Nancy Alvarez as they lead you
through this journey with author Maruja Dagnino.\n\nMiami Beach Botanical GardenÂ 2000
Convention CenterÂ \n\n07:30pm to 11:00pm Red Carpet 07:00pm\n\nÂ \n
event_end_utc 10 Feb 2018 04:00:00
event_name A night of the senses
event_start_utc 10 Feb 2018 00:00:00
facebook_event_id
hashtag None
organization_id la-fontana-ristorante
organization_name La Fontana Ristorante
tags []
venue_city Miami Beach
venue_name Miami Beach Botanical Garden
venue_state FL
venue_street 2000 Convention Center Dr
venue_timezone US/Eastern
venue_zip 33139
Name: cdd7d714-f260-11e7-9897-22000a9b28c5
```

```
event_description Night of the senses; an exclusive gala benefiting the
ArtesanoGroup Foundation. An engaging Dinner and Silent Auction inspired by the
cookbook Los Alimentos del Deseo by the Venezuelan Chef and Journalist Maruja
Dagnino. Join Special guests Boris Izaguirre, George Harris, and Dr. Nancy Alvarez
as they lead you through this journey with author Maruja Dagnino. Miami Beach
Botanical Garden 2000 Convention Center 07:30pm to 11:00pm Red Carpet 07:00pm
event_end_utc 2018-02-10T04:00:00Z
event_name A night of the senses
event_start_utc 2018-02-10T00:00:00Z
facebook_event_id
hashtag None
organization_id la-fontana-ristorante
organization_name La Fontana Ristorante
tags []
venue_city Miami Beach
venue_name Miami Beach Botanical Garden
venue_state FL
venue_street 2000 Convention Center Dr
venue_timezone US/Eastern
venue_zip 33139
Name: cdd7d714-f260-11e7-9897-22000a9b28c5
```



# Phase 2

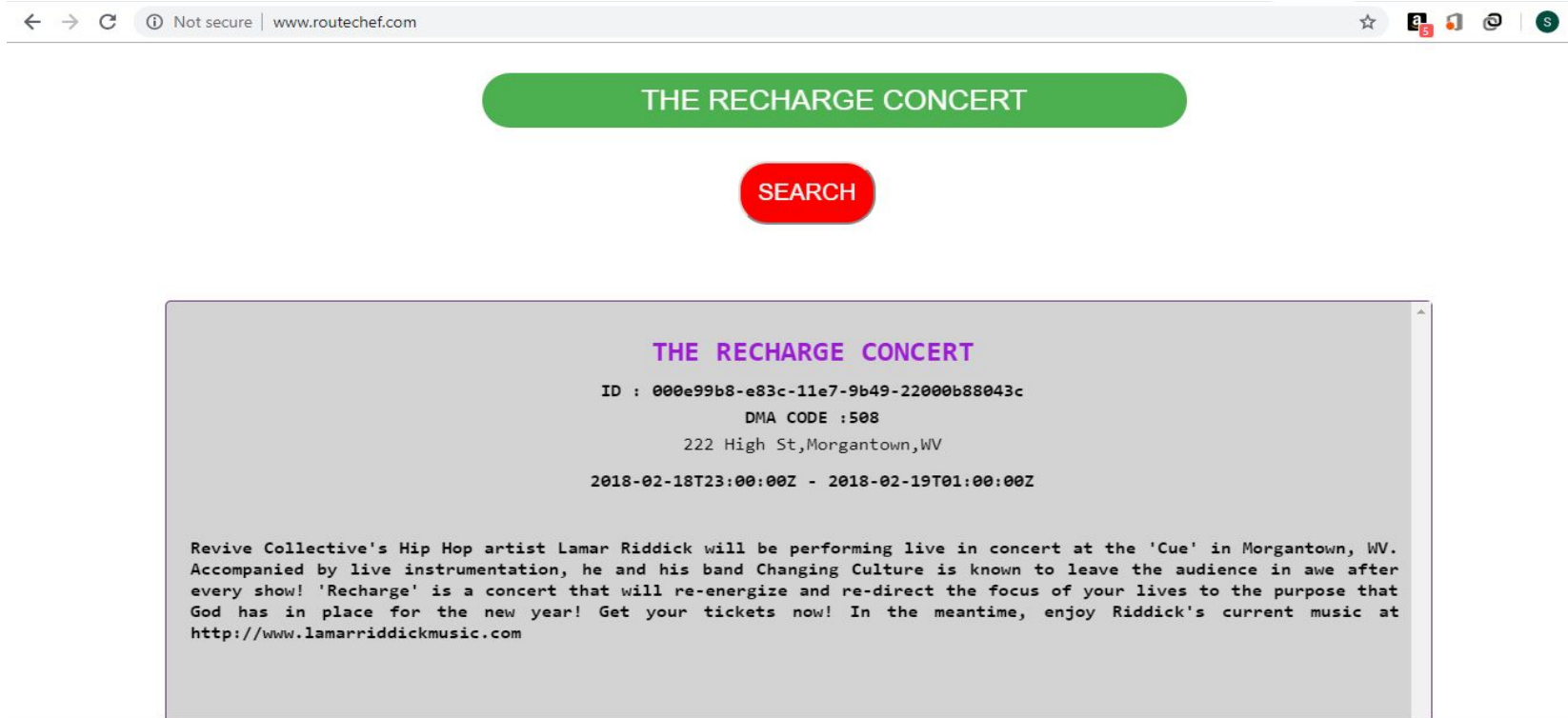
## Display



# DISPLAY ARCHITECTURE

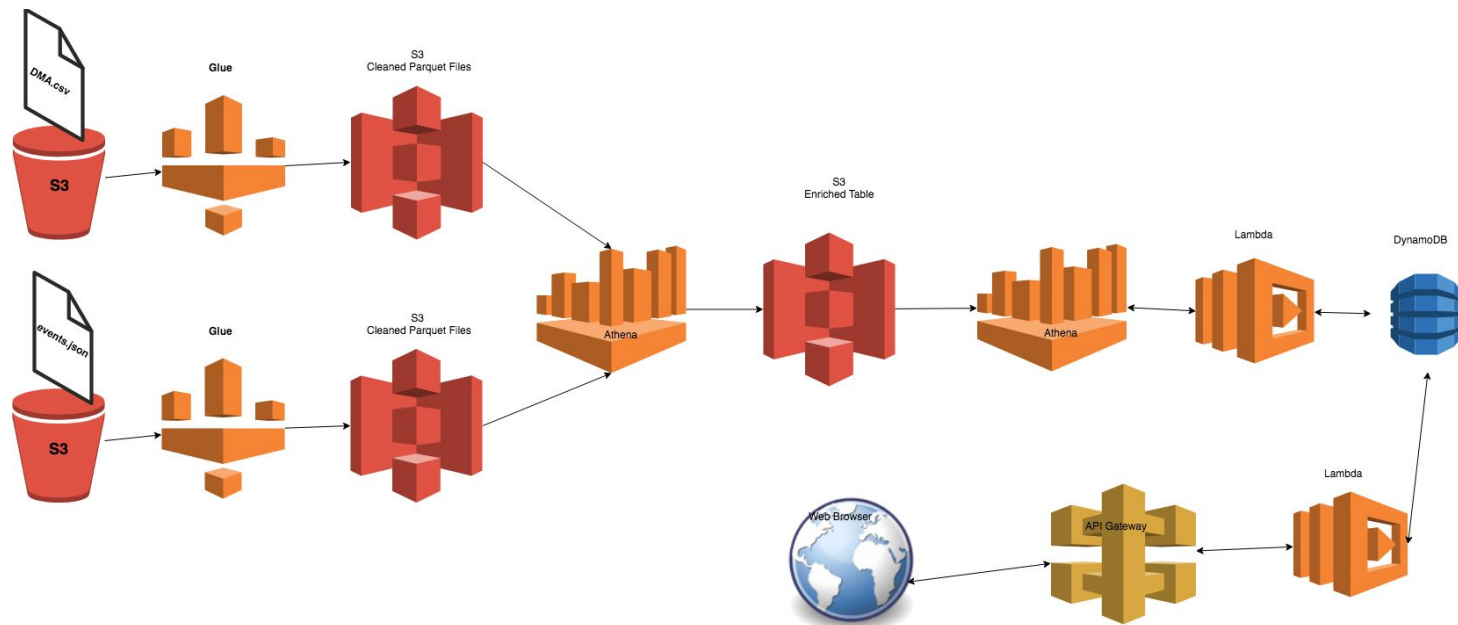


# HTML Page to display event fetched from DYNAMODB



# Phase 3

## Enrichment



event\_description Night of the senses; an exclusive gala benefiting the ArtesanoGroup Foundation. An engaging Dinner and Silent Auction inspired by the cookbook Los Alimentos del Deseo by the Venezuelan Chef and Journalist Maruja Dagnino. Join Special guests Boris Izaguirre, George Harris, and Dr. Nancy Alvarez as they lead you through this journey with author Maruja Dagnino. Miami Beach Botanical Garden 2000 Convention Center 07:30pm to 11:00pm Red Carpet 07:00pm

event\_end\_utc 2018-02-10T04:00:00Z  
event\_name A night of the senses  
event\_start\_utc 2018-02-10T00:00:00Z  
facebook\_event\_id  
hashtag None  
organization\_id la-fontana-ristorante  
organization\_name La Fontana Ristorante  
tags []  
venue\_city Miami Beach  
venue\_name Miami Beach Botanical Garden  
venue\_state FL  
venue\_street 2000 Convention Center Dr  
venue\_timezone US/Eastern  
venue\_zip 33139

Name: cdd7d714-f260-11e7-9897-22000a9b28c5

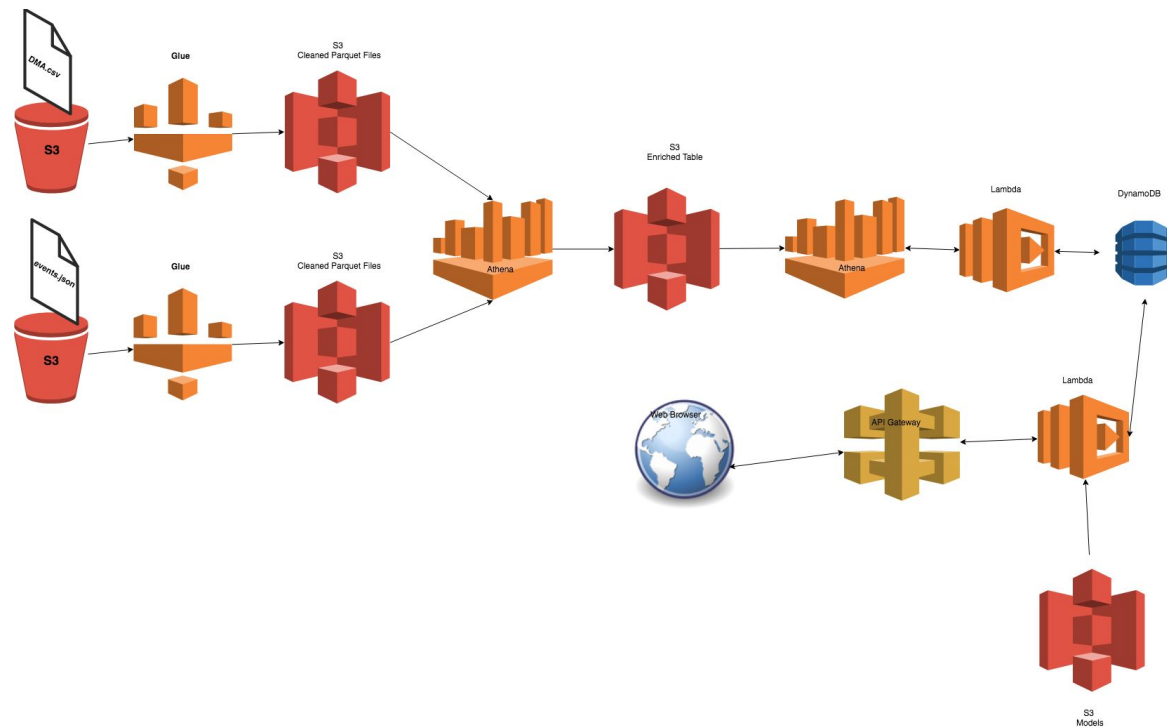
event\_description Night of the senses; an exclusive gala benefiting the ArtesanoGroup Foundation. An engaging Dinner and Silent Auction inspired by the cookbook Los Alimentos del Deseo by the Venezuelan Chef and Journalist Maruja Dagnino. Join Special guests Boris Izaguirre, George Harris, and Dr. Nancy Alvarez as they lead you through this journey with author Maruja Dagnino. Miami Beach Botanical Garden 2000 Convention Center 07:30pm to 11:00pm Red Carpet 07:00pm

event\_end\_utc 2018-02-10T04:00:00Z  
event\_id cdd7d714-f260-11e7-9897-22000a9b28c5  
event\_name A night of the senses  
event\_start\_utc 2018-02-10T00:00:00Z  
facebook\_event\_id NaN  
hashtag NaN  
organization\_id la-fontana-ristorante  
organization\_name La Fontana Ristorante  
tags []  
venue\_city Miami Beach  
venue\_name Miami Beach Botanical Garden  
venue\_state FL  
venue\_street 2000 Convention Center Dr  
venue\_timezone US/Eastern  
venue\_zip 33139  
dma code 528

venue_city	venue_name	venue_state	venue_street	venue_timezone	venue_zip	dma code	dma name
Philadelphia	PHIT Comedy @ The Adrienne Theatre Mainstage	PA	2030 Sansom Street, 1st Floor	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	PHIT Comedy @ The Adrienne Theatre Mainstage	PA	2030 Sansom Street, 1st Floor	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	PHIT Comedy @ The Adrienne Theatre Mainstage	PA	2030 Sansom Street, 1st Floor	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	The Hub at the Cira Center	PA	2001 Market St	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	The Hub at the Cira Center	PA	2001 Market St	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	The Hub at the Cira Center	PA	2001 Market St	US/Eastern	19103	504	PHILADELPHIA
San Diego	Fringe Box Office	CA	969 1st Ave	US/Pacific	92101	825	"SAN DIEGO"
San Diego	Fringe Box Office	CA	969 1st Ave	US/Pacific	92101	825	"SAN DIEGO"
San Diego	Fringe Box Office	CA	969 1st Ave	US/Pacific	92101	825	"SAN DIEGO"
Philadelphia	Philly Improv Theater	PA	2030 Sansom Street	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	Philly Improv Theater	PA	2030 Sansom Street	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	Philly Improv Theater	PA	2030 Sansom Street	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	PHIT Comedy @ The Adrienne Theatre Second Stage	PA	2030 Sansom Street, 2nd Floor	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	PHIT Comedy @ The Adrienne Theatre Second Stage	PA	2030 Sansom Street, 2nd Floor	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	PHIT Comedy @ The Adrienne Theatre Second Stage	PA	2030 Sansom Street, 2nd Floor	US/Eastern	19103	504	PHILADELPHIA
SF	Omni San Francisco Hotel	CA	500 California St	US/Pacific	94104	807	"SAN FRANCISCO-OAKLAND-SAN JOSE"
SF	Omni San Francisco Hotel	CA	500 California St	US/Pacific	94104	807	"SAN FRANCISCO-OAKLAND-SAN JOSE"
SF	Omni San Francisco Hotel	CA	500 California St	US/Pacific	94104	807	"SAN FRANCISCO-OAKLAND-SAN JOSE"
Philadelphia	PHIT Comedy @ The Adrienne Theatre Mainstage	PA	2030 Sansom Street, 1st Floor	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	PHIT Comedy @ The Adrienne Theatre Mainstage	PA	2030 Sansom Street, 1st Floor	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	PHIT Comedy @ The Adrienne Theatre Mainstage	PA	2030 Sansom Street, 1st Floor	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	PHLCVB Office	PA	1601 Market Street, Suite 200	US/Eastern	19103	504	PHILADELPHIA
Philadelphia	PHLCVB Office	PA	1601 Market Street, Suite 200	US/Eastern	19103	504	PHILADELPHIA

# Phase 4

## Analysis



# Our Original Concept

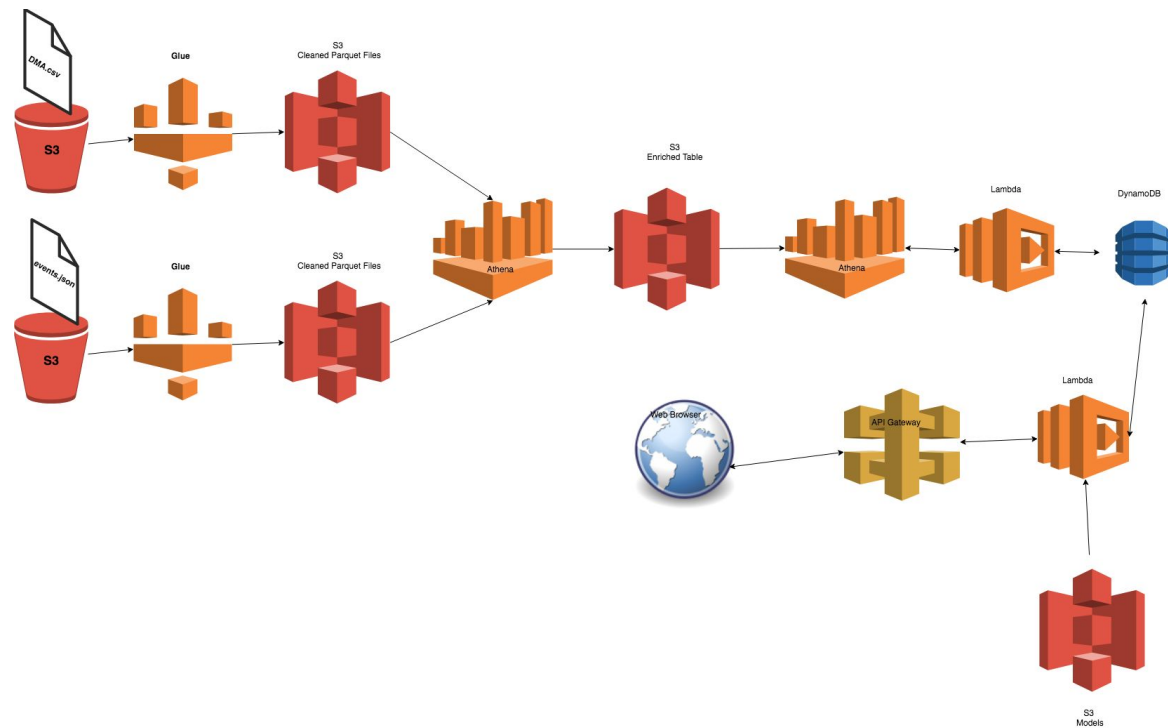


**DMA Code**

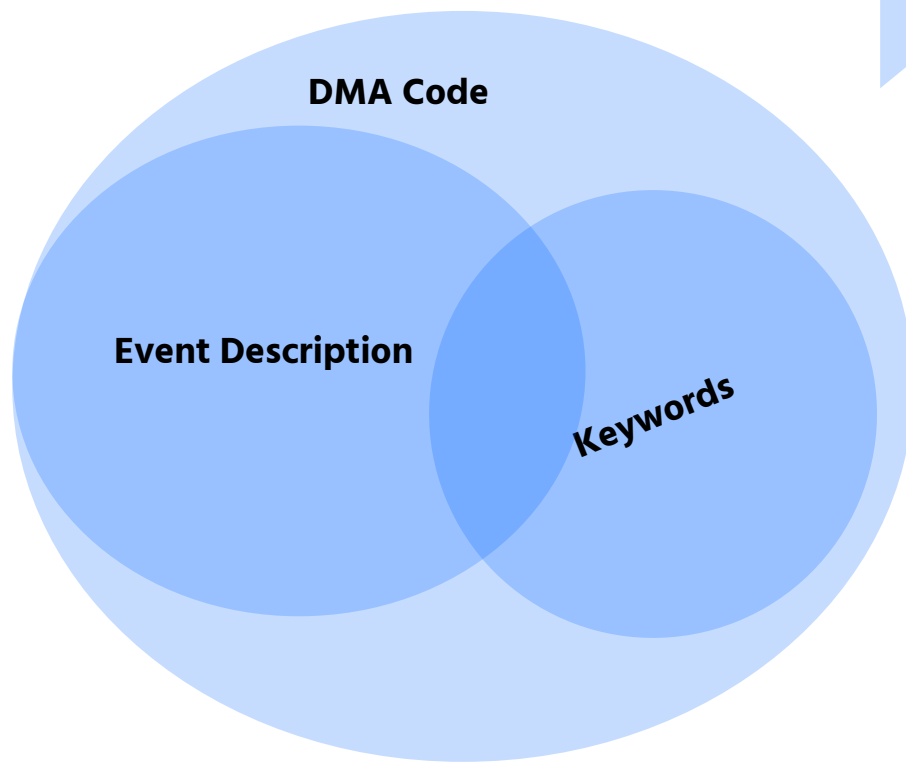


# Phase 5

## Iteration



# Our Concept



# The Original Recommendation Algorithm



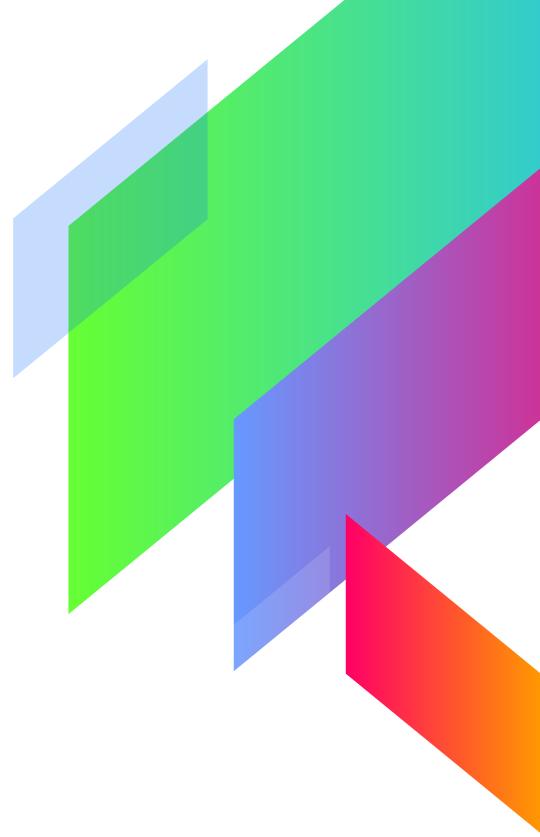
# Pitfalls of the Original Recommendation Algorithm

- Time to obtain recommendations
  - Tfidf model trained with each inquiry
  - Similarity calculated on full dimensionality of Tfidf matrix
- Quality of recommendations
  - Choice of Jaccard similarity
  - No stop words removed



# Optimizing the Algorithm

Using more than Jaccard similarities  
and venue locations for  
recommendations



# Optimizing the Recommendation Algorithm

- Optimize time to obtain recommendations
  - Train models locally and store in S3 for future loading
  - Before pairwise similarity calculations:
    - Reduce dimensionality of Tfidf matrix
    - Cluster events to filter for relevancy
- Improve quality of resulting recommendations
  - Removal of English stop words
  - Change to cosine similarity instead of Jaccard

# Overall Recommendation Scheme

Select only events in the same DMA code



Vectorize event descriptions using Tfidf scores



Reduce dimensionality using singular value decomposition



Cluster using the  $k$ -means algorithm



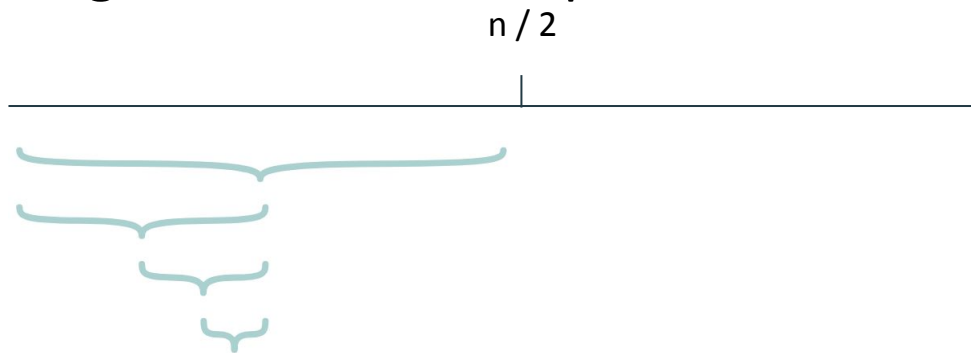
Calculate cosine similarity



Select top 5 recommended events

# Training the Models

- All generated using scikit-learn
- 3 pre-trained models stored in S3 in Python pickle format
  - Tfidf vectorizer
  - Singular value decomposition



- *K*-means clustering





Demo Time!

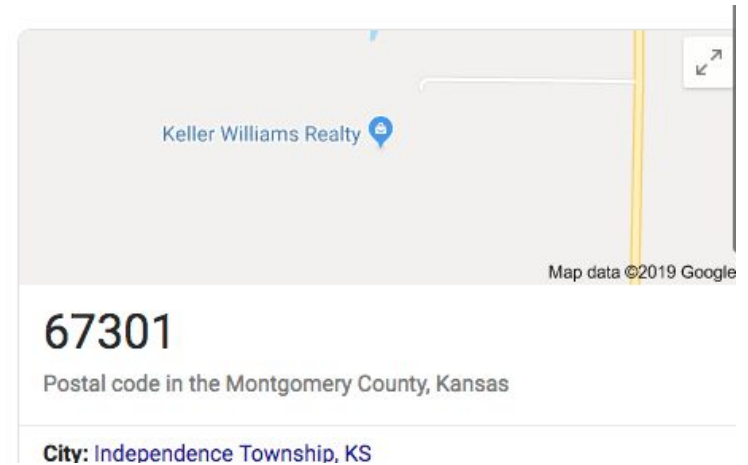
[WWW.ROUTECHF.COM](http://WWW.ROUTECHF.COM)



# Problems We Encountered

- Limited access to AWS resources
- Limited access to high performance computing resources
- Version compatibility issues
- Event Data Accuracy

```
{
  "event_description" : "#MisBeHavenThursdays on April 26th @ Haven Nightclub located inside Golden Nugget Hotel AC features DJ Jason Weiss and Montone!\n\n$5 Corona and Miller Lites\n\n$30 Shot Carafes\n\n$200 Bottles of Patron Tequila and Ultimat Vodka.\n\n$350 Ketel One Magnum Bottles\n\nCall 215.694.8584 for VIP Bottle Service & Discounted Hotel Rooms.\n\nINQUIRE WITHIN FOR GUESTLISTS, VIP BOTTLE SERVICE AND HOTEL ROOMS EVERY THURSDAY, FRIDAY & SATURDAY NIGHT.\u00c2\u00a0=====I Contact Information l215.694.8584Aleckrayzel@yahoo.comwww.twitter.com/AKPROMOSwww.instagram.com/akpromotionshttp://www.facebook.com/akpromotionshttp://www.facebook.com/groups/AKPROMOTIONSwww.AK-Promotions.com\n",
  "event_end_utc" : "27 Apr 2018 08:00:00",
  "event_id" : "7361f56c-3c6b-11e8-abd7-22000b6f049c",
  "event_name" : "Jason Weiss & Montone @ Haven Nightclub AC Thursday April 26th",
  "event_start_utc" : "27 Apr 2018 03:00:00",
  "facebook_event_id" : "187689625185208",
  "hashtag" : null,
  "organization_id" : "akpromotions",
  "organization_name" : "AK PROMOTIONS",
  "tags" : [],
  "venue_city" : "Atlantic City",
  "venue_name" : "HAVEN NIGHTCLUB",
  "venue_state" : "NJ",
  "venue_street" : "1 Castle Blvd",
  "venue_timezone" : "US/Eastern",
  "venue_zip" : "67301"
},
```



## Ways to Further Improve

- Include additional stop words (“ticket,” prices, days of the week)
- Improve time to get recommendations
  - Pre-calculate recommendations and store as a column in Athena/DynamoDB
  - Use locality-sensitive hashing instead of pairwise similarity calculations
- Incorporate date theory
- Improve ability of algorithm to identify nearly duplicate events
- Incorporate more information (event name, hashtags, etc.)
- Check Validity of Event (ex. Venue Zip Code and Venue State)

# Skills that We Used

- Term Frequency-Inverse Document Frequency
  - Similarity Metrics
- Singular Value Decomposition
  - *K*-Means Clustering

Abstract geometric shapes in the bottom-left corner, including a green triangle, a light blue parallelogram, a brown trapezoid, a red parallelogram, and a purple parallelogram.

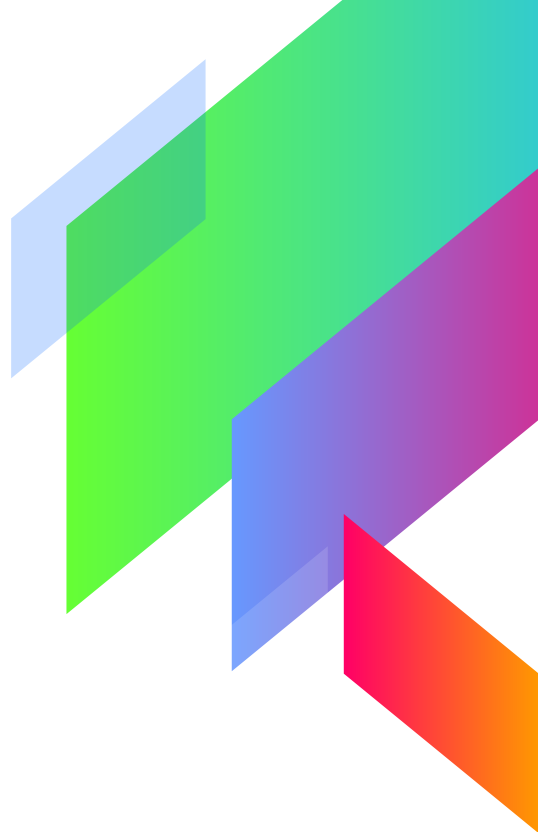
**Thank You,  
Mr. Young!**

Abstract geometric shapes in the bottom-right corner, including a light blue parallelogram, a green triangle, a purple parallelogram, a red parallelogram, and an orange parallelogram.



# THANKS!

**Any questions?**



GitHub Link: <https://github.com/heatherciallella/recommendation-engine>