

# POIREVIEWQA: A SEMANTICALLY ENRICHED POI RETRIEVAL AND QUESTION ANSWERING DATASET

*GIR'18, Nov. 2018*



**Gengchen Mai<sup>1</sup>** Krzysztof Janowicz<sup>1</sup>  
**Cheng He<sup>2</sup>** Sumang Liu<sup>2</sup> Ni Lao<sup>2</sup>

<sup>1</sup> STKO Lab, University of California Santa Barbara  
<sup>2</sup>SayMosaic Inc., Palo Alto, CA, USA

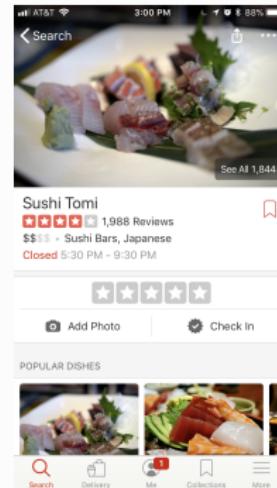


# LOCATION-BASED SERVICES & POI SEARCHING

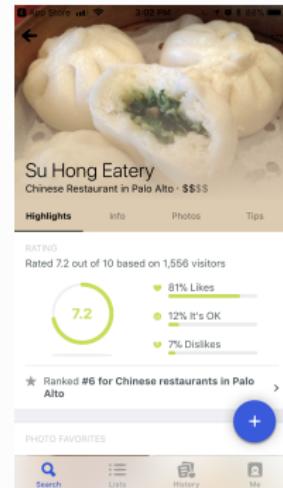
- **Location-based services (LBS)** and the underlying **Point of Interest (POI)** datasets play a increasingly important role in our daily interaction with mobile devices.



Google Place



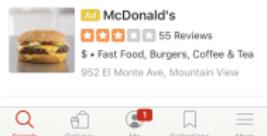
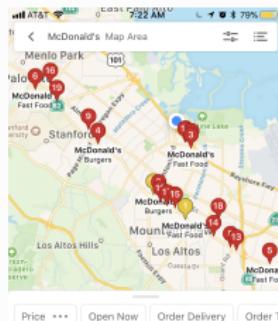
Yelp



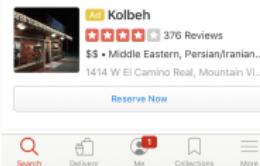
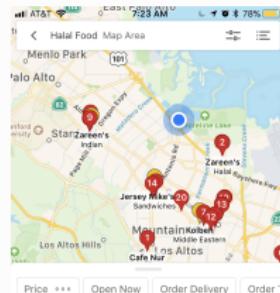
Foursquare

# LOCATION-BASED SERVICES & POI SEARCHING

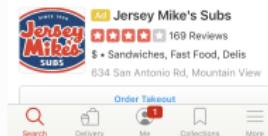
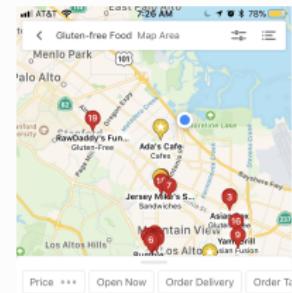
- Platforms such as Yelp, Foursquare, Google Map allow users to search nearby POIs based on their **names**, **place types**, or **tags**, which requires manual data annotation.



Place Name Search



Place Type Search



Tag Search

# PROBLEMS IN CURRENT POI SEARCHING

- **Predefined Tags:** The tags (take reservations, delivery, take out) are also **predefined** and **manually added** which may not satisfy users' needs and require a lot of manually efforts.

## Hours

Mon	5:30 pm - 9:30 pm
Tue	5:30 pm - 9:30 pm
Wed	5:30 pm - 9:30 pm <b>Closed now</b>
Thu	5:30 pm - 10:00 pm
Fri	5:30 pm - 10:00 pm
Sat	5:30 pm - 10:00 pm
Sun	Closed

[Edit business info](#)

## Upcoming Special Hours

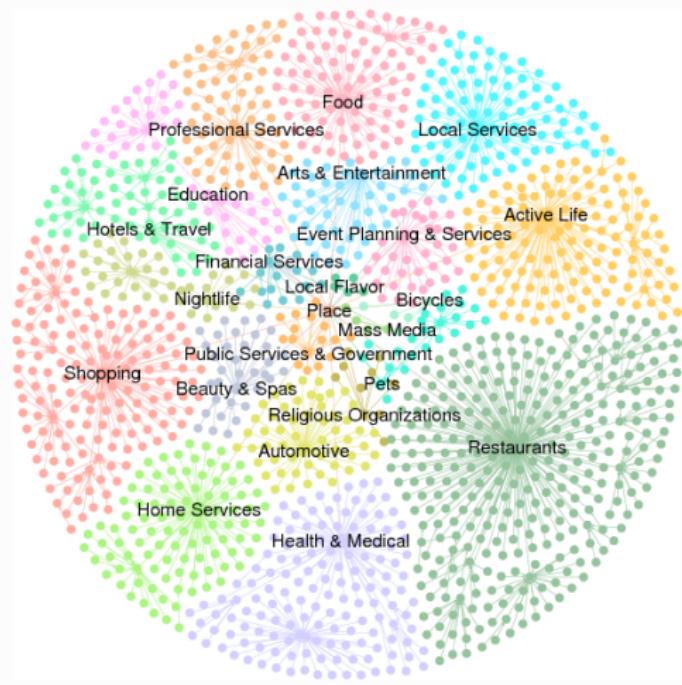
Thu 11/22/2018	Closed
Mon 12/24/2018	Closed
Mon 12/31/2018	Closed

## More business info

Liked by Vegetarians	Yes
Takes Reservations	Yes
Delivery	No
Take-out	No
Accepts Credit Cards	Yes
Good for Kids	No
Good for Groups	Yes
Outdoor Seating	Yes

# PROBLEMS IN CURRENT POI SEARCHING

- **Predefined Schema:** The place type schema is **predefined** and unbalanced.



## INTRODUCTION

- Do we utilize all info in Yelp for search?
- Yelp contains billions of Points of Interest (POIs) and their attributes:
  - name
  - location
  - type
  - open times
  - rating
  - **reviews**
  - ...
- Can we use other **crowdsourcing information** e.g. reviews to **enrich POI** and facilitate POI searching process?

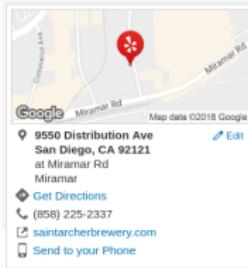


# ANSWERING QUESTIONS BY POI REVIEWS

## Saint Archer Brewery Claimed

★★★★★ 177 reviews [Details](#)

\$ - Breweries, Beer, Wine & Spirits [Edit](#)



[Write a Review](#)

[Add Photo](#) [Share](#) [Save](#)



Photo of Saint Archer Brewery - San Diego, CA, United States

[See all 338](#)

## Ask the Community

### Is this place dog friendly?

 They have an outdoor patio and I've seen customers with dogs sitting outside.  
11 months ago • 1 person found this helpful

[View 1 more answer](#)

★★★★★ 12/3/2017

Awesome brewery. There are so many breweries in San Diego but this one definitely stands out. It's fairly inexpensive and it's **dog friendly**. Both indoor and outdoor seating areas, you know since it's sunny in San Diego year round.

As far as the beer goes, what's your speed? From stouts to IPAs, they have you covered.



# ANSWERING QUESTIONS BY POI REVIEWS

## Fufu's Mid East Grill And Hookah

Claimed

★★★★★ 324 reviews [View Details](#)

\$S - Mediterranean, Hookah Bars, Middle Eastern



195 E Alessandro Blvd  
Ste 8D  
Riverside, CA 92508  
[Get Directions](#)  
(951) 789-0004  
[fufusgrill.com](#)  
[Message the business](#)  
[Send to your Phone](#)



[Write a Review](#)

[Add Photo](#)

[Share](#)

[Save](#)

[See all 216](#)

★★★★★ 6/1/2018

4 check-ins

Really, really like this place. Have been here a few times. Probably the only Mediterranean/Middle Eastern food in Riverside and the food is great. Healthy, tasty, taste fresh, service is excellent, good price and a very good variety on the menu.

The food is **halal** too, so for the Muslim community this is THE place to go to!

Only negative is that there is no WiFi, so I just about checked checked in on Yelp here.

[Mark B. and 1 other](#) voted for this review

Useful 2

Funny

Cool 2

## Ask the Community

Hello, do you serve **HALAL** meat? If so, do you have the certification that shows proof? Thank you!



Sarah B. of Fufu's Mid East Grill And Hookah  
Business Owner

Yes everything we serve is 100% **halal**. And we do have the certificate of proof at the restaurant  
Thanks

2 months ago • 3 people found this helpful

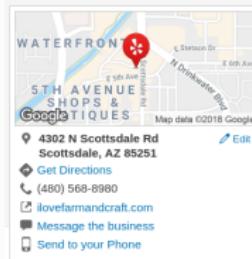
[View question details](#)

# ANSWERING QUESTIONS BY POI REVIEWS

## Farm & Craft

 541 reviews [View Details](#)

\$\$ - American (New), Breakfast & Brunch, Cocktail Bars [Edit](#)



 Write a Review

 Add Photo  Share  Save



See all 636

## Ask the Community

### Dog friendly patio?

 Yes!! Saw lots of cute dogs sitting in the patio  
5 months ago • 2 people found this helpful

[View 5 more answers](#)

### Is your salmon poke bowl **gluten free** and dairy free?

 Kevin B. of Farm & Craft  
Business Customer Service

Hey Bailee, the dish is dairy free, the dish is not **gluten free** because the salmon marinade has a slight amount of gluten in it from the soy sauce. However, upon request we can use **gluten free** soy sauce for the salmon but it may take a bit longer to... [Read more](#)

 5/16/2018

 1 check-in

Food is good, cheap, **gluten free**

Defiantly more about drinks

Tried to get them to serve me dinner at 8:40pm - no go.... not impressed, bet the owner wouldn't be either

Katie T. voted for this review

 Useful 1  Funny  Cool

# POI REVIEW IR & QA

- **A Challenging Task: semantically enriching POIs from unstructured data**, e.g. POI reviews, in order to support open-domain **search** and **question answering (QA)**.
- **One Approach**: Given a question, e.g. **find night clubs near me which are 18+**, build a model to:
  - **Generate the answer** to the question given the current POI review
  - **Rank POI reviews** based on whether they can answer it and the distance from POIs to the user location.
- Problem: **No existing benchmark dataset to train and evaluate such kind of models**
  - **No GIR Corpus**: Only a few existing GIR benchmark datasets (e.g., GeoCLEF); lack in rich annotations.
  - **No QA Corpus**: Many QA datasets for reading comprehension such as SQuAD; no spatial/platial component; **very different query sets**.

## OUR CONTRIBUTION

- We introduce a evaluation benchmark dataset **POIRReviewQA** which consists of **20k questions** for **1022 Yelp business types** to support open-domain POI search and QA.
- We build two baseline models: a **Lucene-based baseline model** and a **sentence embedding-based model** which indicate that the dataset presents a challenging problem for future research by the GIR community

# POIREVIEWQA

- We created POIRReviewQA based on the **Yelp Challenge 11 (YC11) dataset** and the “Ask the Community” section of POI pages.
- Yelp platform is **dominated** by popular business types such as restaurants.
- We need to produce a **balanced query set** for all business types.

The screenshot shows the homepage of the Yelp Dataset Challenge. At the top, there's a red header bar with the "yelp Dataset" logo and links for "Dataset", "Challenge", and "Documentation". Below the header, a large banner features a pink illustration of a town with buildings, mountains, and a clock tower. The text "Yelp Dataset Challenge" and "Discover what insights lie hidden in our data." is displayed. A section titled "What is the dataset challenge?" explains the purpose of the challenge, mentioning research, analysis, and cash prizes. Another section, "The Challenge", lists topics like Photo Classification and Graph Mining, each accompanied by small images related to the topic.

# QUERY SET GENERATION

- Count the **frequencies of POI name suffixes** (single words) in YC11
- For every suffix ( $\text{freq} \geq 10$ ), we create a **quoted search query**, and collect QA page URLs from **Google search engine**:
  - `site:yelp.com/questions/ "Restaurant Ask the Community"`
  - **1,701** quoted search queries
  - **Up to 100** search results for each query
- Collect questions and answers from the **community QA pages**

The screenshot shows a Yelp search results page for "Majoria's Commerce Restaurant". A user named Grety H. asked a question: "Is this place breakfast all day?". Below the question, there are 6 answers. The first answer is from Jim S., who says "More or less, but they close about 230-300 pm". The second answer is from Monica T., who says "they start to run out of breakfast items around 11 so I would go there early for breakfast!". The third answer is from Joanna C., who says "This place serves breakfast until it closes at 2:30PM". The fourth answer is from Steffan O., who says "yes, get there early, great food, great prices, great service, you will love it.". The fifth answer is from Lisa G., who has not provided an answer.

## QUERY SET GENERATION

- Since the suffix often indicates the place type of a POI, the collected query set has **a wide coverage of different place types**.
- **20K questions** were collected from Yelp business question pages for **1022 Yelp business types**.
- Each question is associated with one or multiple POIs with several POI types.

Question: **Age to get in?**

Context:

- [Echoplex](#) [Music Venues, Bars, Dance Clubs]
- [Paper Tiger Bar](#) [Cocktail Bars, Lounges]

## RELEVANCE AND ANSWER ANNOTATION

- For each question, **10 review candidates** are selected:
  - Apply **Elastic Search** to POI reviews based on the question with **constraint to the associated POI types**
  - Do **stratified sampling** on the search result



elasticsearch

# RELEVANCE AND ANSWER ANNOTATION

Any gluten-free food available?

Search

Here shows some search options:

Pseudo Revelance Feedback

[Return 1000 results](#)

**0. Brix Wood Fired Pizza**  
Restaurants, Pizza  
**21.77325**

This is still my goto choice for **gluten free** pizza. The have **gluten free** beer as well. The appetizer selections that are **gluten free** are pretty limited. They have no **gluten free** desserts at all. But the pizza selection is fabulous. My top three are the rustica, the greek, and the chicken and mushroom. If you com at 10:00 or later on Sunday through Thursday the pizza is 2 for 1. My wife likes the sandwiches and salads. Kids are welcome but there is no mac and cheese. Nor are there any burgers available. Enjoy.

**1. Humble Pie**  
Italian, Pizza, Restaurants  
**21.35636**

Very sad that **gluten free** crust any available topping. Crust needs something more to be pizza. Had to leave to get dinner.

**2. The Cheesecake Factory**  
American (New), American (Traditional), Food, Desserts, Restaurants  
**21.21855**

We always look forward to coming up to the Cheesecake Factory. They offer quite the variety of food. Also most all their pastas are available in a **gluten free** option, and tastes great. Sad to say most of the time I'm to stuffed to make it to their cheesecake. Which I think is some sort of foodie code violation. But you can get a piece to go or better yet a whole cake frozen to go. Just make sure when you order **gluten free** to remind the kitchen to make any meat toppings **gluten free** also. (Like grilled chicken instead of breaded)

**3. Famous 48**  
American (New), Restaurants  
**20.990713**

Great for **gluten free** menu! We had the deviled eggs, spinach con queso and bbq baby back ribs. Great food, service and atmosphere. **Gluten free** pizza crust and buns available!

**4. PPG Paints Arena**  
Stadiums & Arenas, Event Planning & Services, Venues & Event Spaces, Music Venues, Nightlife, Arts & Entertainment  
**20.95517**

5 stars for layout, cleanliness, friendly staff, and overall quality for a venue. 1 star for **gluten free** options. There are close to zero **gluten/allergen** free items in this entire arena, save the fruit options at some concession stands. The nachos are **gluten free** but processed with other wheat products and also fried in soybean oil, the salad has granola mixed in with it, and the sushi IS **gluten free**, but no safe soy sauce is available. They sell Redbridge lager, which is great, but I found myself having beer for dinner because I couldn't eat anything. They will also

**5. Make It Yourself Pizza**  
Gluten-Free, Restaurants, Pizza  
**20.874434**

Awesome fresh baked pizza. No more reheated slices. They'll put any amount of toppings on it all for one price! **Gluten free** crust is available too!!

**6. Cafe Monte**  
Restaurants, Food, Breakfast & Brunch, French, Bakeries  
**20.846643**

**7. Cafe Rio**  
Mexican, American (Traditional), Salad, Restaurants  
**20.7883**

**8. Giovanni's Pizza & Pasta**  
Restaurants, Gluten-Free, Pizza  
**20.779844**

## RELEVANCE AND ANSWER ANNOTATION

- We developed a **crowd-facing Web server** and deployed it on **Amazon Mechanical Turk** to let raters annotate the answer for each question given POI reviews.
- Each rater is given **a question** and **10 review candidates** associated with it which are divided into **sentences**.
- For each sentence:
  - **Whether it answer the current question**
  - **What the corresponding answer is**
- Each question is annotated by **3 different raters** to avoid subject biases.

# RELEVANCE AND ANSWER ANNOTATION

## Survey Instructions (Click to collapse)

### Help Label Yelp Business Reviews

This task aims at choosing which Yelp review sentences answer the given business question.

- For this task, we are giving you a question which was posted on Yelp regarding a specific business (**Source Business**), i.e., **Age to get in?**
- Source Businesses and Types** are provided to let you have a better understanding of the context of this question, i.e., [Echoplex](#) [Music Venues & Bars, Dance Clubs];
- 10 Yelp reviews from different businesses are provided. Each review is divided into sentences;
- Please select one or multiple sentences which answer the question asked, i.e., **Wide variety of people of all ages 21+**;
- Please also summarize the answer(s) you chosen in your own words, i.e., **21 or older**;
- Note: if none of the sentences is correct, please **ONLY** select **None of them** and leave the summarization text box **EMPTY**;
- Note: these 10 reviews are not related to each other. Each review comes from a different business, which may provide different answers. All them are considered correct, i.e. **Review A: 21 or older or Review B: 20 or older**

An Example Question and Review Pair:

Question: **Age to get in?**

Context:

- [Echoplex](#) [Music Venues, Bars, Dance Clubs]
- [Paper Tiger Bar](#) [Cocktail Bars, Lounges]

**Review 1. [Source Business]**

1. Please select the sentence(s) from this Yelp business review which answer(s) the current question:

- Excellent place to country dance and make new friends.
- Drinks are reasonably priced, especially domestic beer before 10pm on certain nights.
- Wide variety of people of all ages 21+.
- This is my favorite country bar in the east valley.
- None of them

2. If you find (a) sentence(s) which answers the question, please summarize the answer and enter it below:

21 or older

# RELEVANCE AND ANSWER ANNOTATION

Survey Instructions (Click to expand)

Question: Am I able to bring in my own bottle?

Context:

- Waterman Grille [American (New), Seafood]

Review 1. [Source Business]

1. Which of the following answer the above question:

- Absolute wonderful place!
- My family was in town for a wake/funeral and due traffic did not have time for dinner before the viewing.
- We called ahead to make sure they would be able to accommodate our party later in the evening.
- We also mentioned that we had Campari lovers and the owner went out and got a bottle for us!
- Our server was delightful and the food was outstanding!
- I ordered the steak and frites and it was delicious.
- Everyone at the table ordered something different and each person loved their meal!
- Next time we are in town we will DEFINITELY be back!
- None of them

2. Please summarize your selection with a few words:

Enter your answer here...

Review 2. [Source Business]

1. Which of the following answer the above question:

- Great wine and great customer service!
- We were in Scottsdale for my friend's bachelorette party, and we did the "bottling party" at Su Vino.
- We had a big table that sat all 8 of us, with cheese, fruit and crackers, and our own personal wine maker bringing us tastes and explaining the different wines.
- Then we were able to bottle our own wine, and put our own label on it (we had a picture of all of us girls).
- It was a blast!

I would definitely recommend a visit here!

## RELEVANCE AND ANSWER ANNOTATION

- **4100** out of 20K questions have been annotated until now.

**Table 1.** The Statistic of POIRreviewQA

# of Annotated question	4,100
% of questions WITHOUT related reviews	11.4%
Avg. # of related <i>reviews</i> per question	4.61
Avg. # of 1 rater agreeing on relevant <i>sentence</i> per question	2.19
Avg. # of 2 raters agreeing on relevant <i>sentence</i> per question	1.08
Avg. # of 3 raters agreeing on relevant <i>sentence</i> per question	0.83

## RELEVANCE AND ANSWER ANNOTATION

- For these 4100 questions, **11.4%** of them are **without** any related reviews.

**Table 1.** The Statistic of POIReviewQA

# of Annotated question	4,100
% of questions WITHOUT related reviews	11.4%
Avg. # of related <i>reviews</i> per question	4.61
Avg. # of 1 rater agreeing on relevant <i>sentence</i> per question	2.19
Avg. # of 2 raters agreeing on relevant <i>sentence</i> per question	1.08
Avg. # of 3 raters agreeing on relevant <i>sentence</i> per question	0.83

## RELEVANCE AND ANSWER ANNOTATION

- For these 4100 questions, **4.61** out of 10 reviews are related per question.

**Table 1.** The Statistic of POIReviewQA

# of Annotated question	4,100
% of questions WITHOUT related reviews	11.4%
Avg. # of related <i>reviews</i> per question	4.61
Avg. # of 1 rater agreeing on relevant <i>sentence</i> per question	2.19
Avg. # of 2 raters agreeing on relevant <i>sentence</i> per question	1.08
Avg. # of 3 raters agreeing on relevant <i>sentence</i> per question	0.83

## RELEVANCE AND ANSWER ANNOTATION

- **Interrater agreement** (given 3 raters per review sentence):

- Divide the sentences into 4 sets based on the number of raters that agreed on each sentence:  $R_0$ ,  $R_1$ ,  $R_2$ ,  $R_3$
- **Randomly sample 20 sentences** from each of the last three sets ( $R_1$ ,  $R_2$ ,  $R_3$ )
- **Manually inspect** the relevance of these sentences to the corresponding questions and calculate the accuracy for each set

Acc. of $R_1$	45%
Acc. of $R_2$	90%
Acc. of $R_3$	100%

## RELEVANCE AND ANSWER ANNOTATION

- The average number of sentences from each set ( $R_1$ ,  $R_2$ ,  $R_3$ ) per question

**Table 1.** The Statistic of POIReviewQA

# of Annotated question	4,100
% of questions WITHOUT related reviews	11.4%
Avg. # of related <i>reviews</i> per question	4.61
Avg. # of 1 rater agreeing on relevant <i>sentence</i> per question	2.19
Avg. # of 2 raters agreeing on relevant <i>sentence</i> per question	1.08
Avg. # of 3 raters agreeing on relevant <i>sentence</i> per question	0.83

## EVALUATION METRICS

- Mean average precision (MAP)
- Area Under receiver operating characteristic Curve (AUC)

$$\text{MAP} = \frac{\sum_{q=1}^Q \text{AveP}(q)}{Q}$$

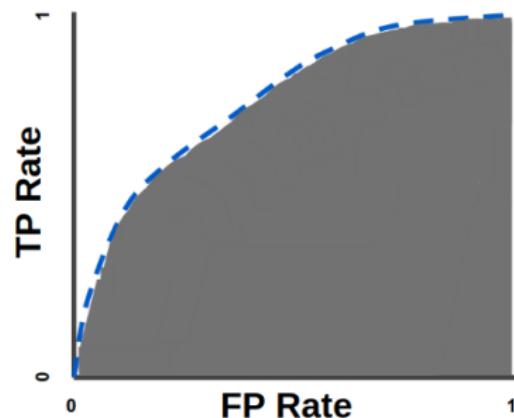
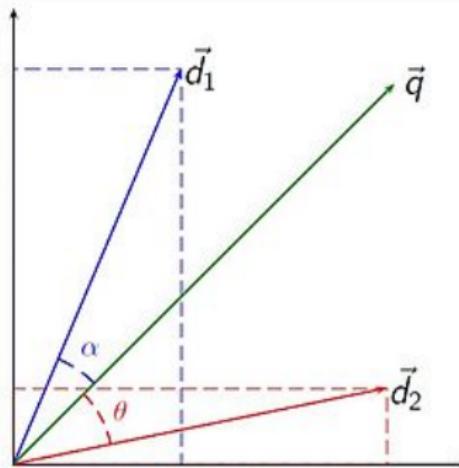


Figure 5. AUC (Area under the ROC Curve).

## BASELINE MODEL

- **Lucene based model** : For each question, we rerank all the sentences from the 10 reviews selected by Elasticsearch by **TF-IDF** model.



# BASELINE MODEL

## ■ Sentence embedding model (Sanjeev Arora et al. ICLR 2017)

---

**Algorithm 1** Sentence Embedding

---

**Input:** Word embeddings  $\{v_w : w \in \mathcal{V}\}$ , a set of sentences  $\mathcal{S}$ , parameter  $a$  and estimated probabilities  $\{p(w) : w \in \mathcal{V}\}$  of the words.

**Output:** Sentence embeddings  $\{v_s : s \in \mathcal{S}\}$

- 1: **for all** sentence  $s$  in  $\mathcal{S}$  **do**
  - 2:    $v_s \leftarrow \frac{1}{|s|} \sum_{w \in s} \frac{a}{a + p(w)} v_w$
  - 3: **end for**
  - 4: Form a matrix  $X$  whose columns are  $\{v_s : s \in \mathcal{S}\}$ , and let  $u$  be its first singular vector
  - 5: **for all** sentence  $s$  in  $\mathcal{S}$  **do**
  - 6:    $v_s \leftarrow v_s - uu^\top v_s$
  - 7: **end for**
-

## EVALUATION

- We evaluate two baseline models against POIRReviewQA by using AUC, MAP

	Lucene Based	Sentence Embedding
AUC	77.0%	79.2%
MAP	48.8%	41.0%

# EVALUATION

- **Q:** Question
- **T:** POI Type
- **C:** Review Sentence
- **A:** Answer
- **H:** Human judgements
- **L:** Lucene
- **E:** Sentence Embedding

Reason	Example	Ranking (H/L/E)
Paraphrase	<p><b>Q:</b> About <b>how long</b> should I expect my visit to be?</p> <p><b>T:</b> Venues &amp; Event Spaces; Kids Activities</p> <p><b>C:</b> We <b>were there for</b> about 2 hours, including the show.</p> <p><b>A:</b> took 2 hrs</p>	1/107/88 out of 158
Hyponym	<p><b>Q:</b> Any good <b>vegan choices</b>?</p> <p><b>T:</b> Restaurants→Cajun/Creole</p> <p><b>Sent:</b> After scanning the menu for a bit however, I was able to find the <b>tofu</b> wings.</p> <p><b>A:</b> Tofu wings could be a choice</p>	2/49/18 out of 83
Synonymy	<p><b>Q:</b> Any recommendations on how to <b>score a table</b>? ...</p> <p><b>T:</b> Restaurants→French</p> <p><b>C:</b> I <b>made a reservation</b> a day in advance thinking it will be busy.</p> <p><b>A:</b> A day in advance</p>	1/63/14 out of 98
Deduction	<p><b>Q:</b> Are there classes for <b>seniors</b>?</p> <p><b>T:</b> Art Galleries; Art Schools</p> <p><b>C:</b> Great studio for <b>all</b>, including kids!</p> <p><b>A:</b> There are classes for seniors</p>	1/60/15 out of 72
Common Sense	<p><b>Q:</b> Do they <b>buy comic books</b>?</p> <p><b>T:</b> Shopping→Comic Books</p> <p><b>C:</b> The concerns: The store currently has <b>no consignment or new issues</b>.</p> <p><b>A:</b> No</p>	1/45/53 out of 62

## CONCLUSION

- we introduce a evaluation benchmark dataset **POIReviewQA**<sup>1</sup> which consists of 20k questions for 1022 Yelp business types to support open-domain POI search and QA.
- We build two baseline models: a **Lucene-based baseline model** and a **sentence embedding-based model** which indicate that the dataset presents a challenging problem for future research by the GIR community
- The POIReviewQA Task follows the idea of **semantic signatures** for social sensing by which we can study POI types using patterns extracted from human behavior, e.g., what people write about places of a particular type.

---

<sup>1</sup><http://stko.geog.ucsb.edu/poireviewqa/>