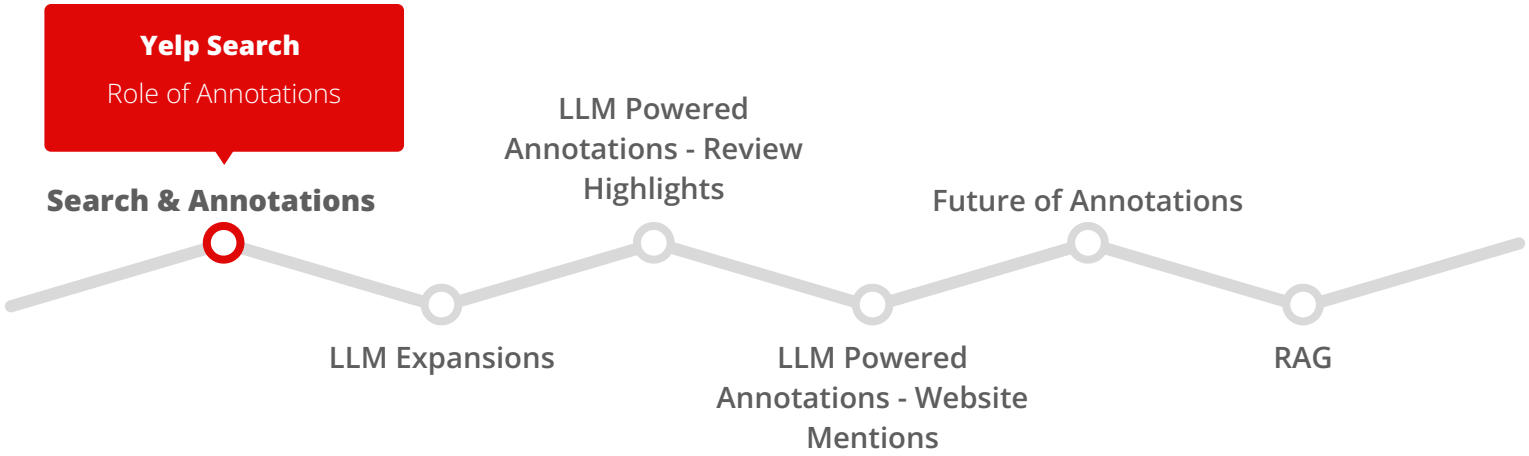
RELEVANCE PROOF IN YELP SEARCH: LLM-POWERED ANNOTATIONS

Pallavi Patil

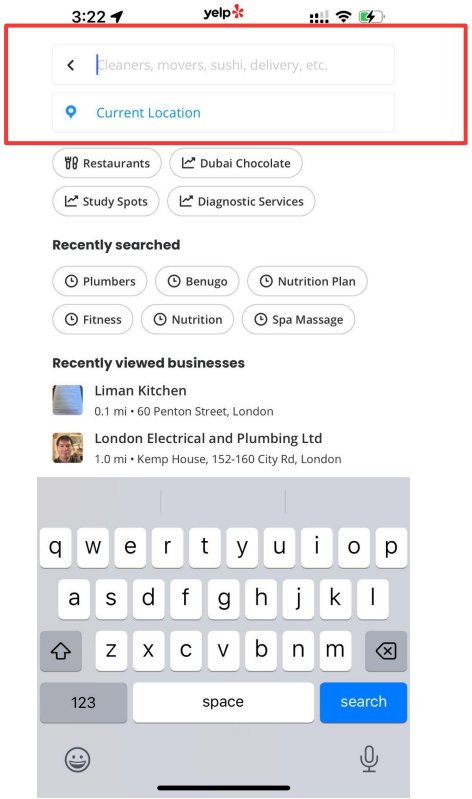
yelp


Yelp Search
Role of Annotations
Search & Annotations
LLM Expansions
LLM Powered Annotations - Review Highlights
LLM Powered Annotations - Website Mentions
Future of Annotations
RAG

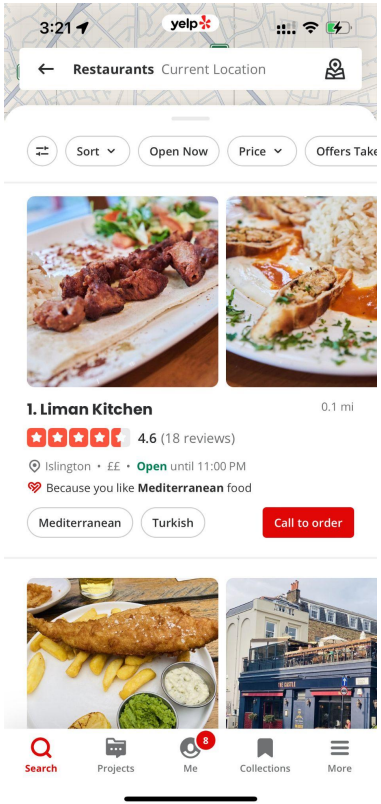
Yelp Search

yelp
tacos, cheap dinner, Max's
New York, NY 10001
Yelp for Business
Write a Review
Restaurants
Home Services
Auto Services
More
Yelp connects people with great local businesses.

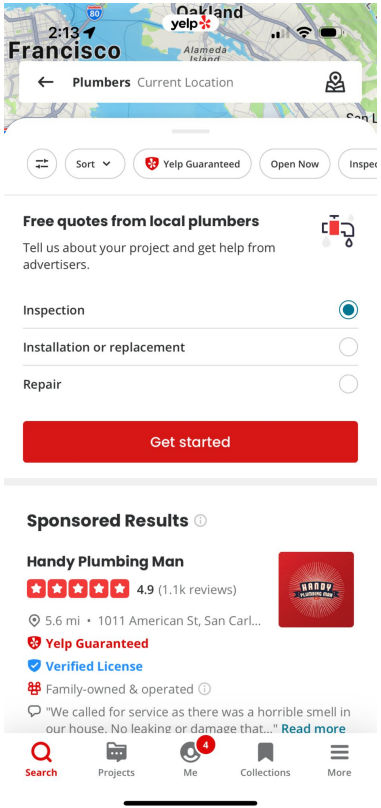
Search Flow



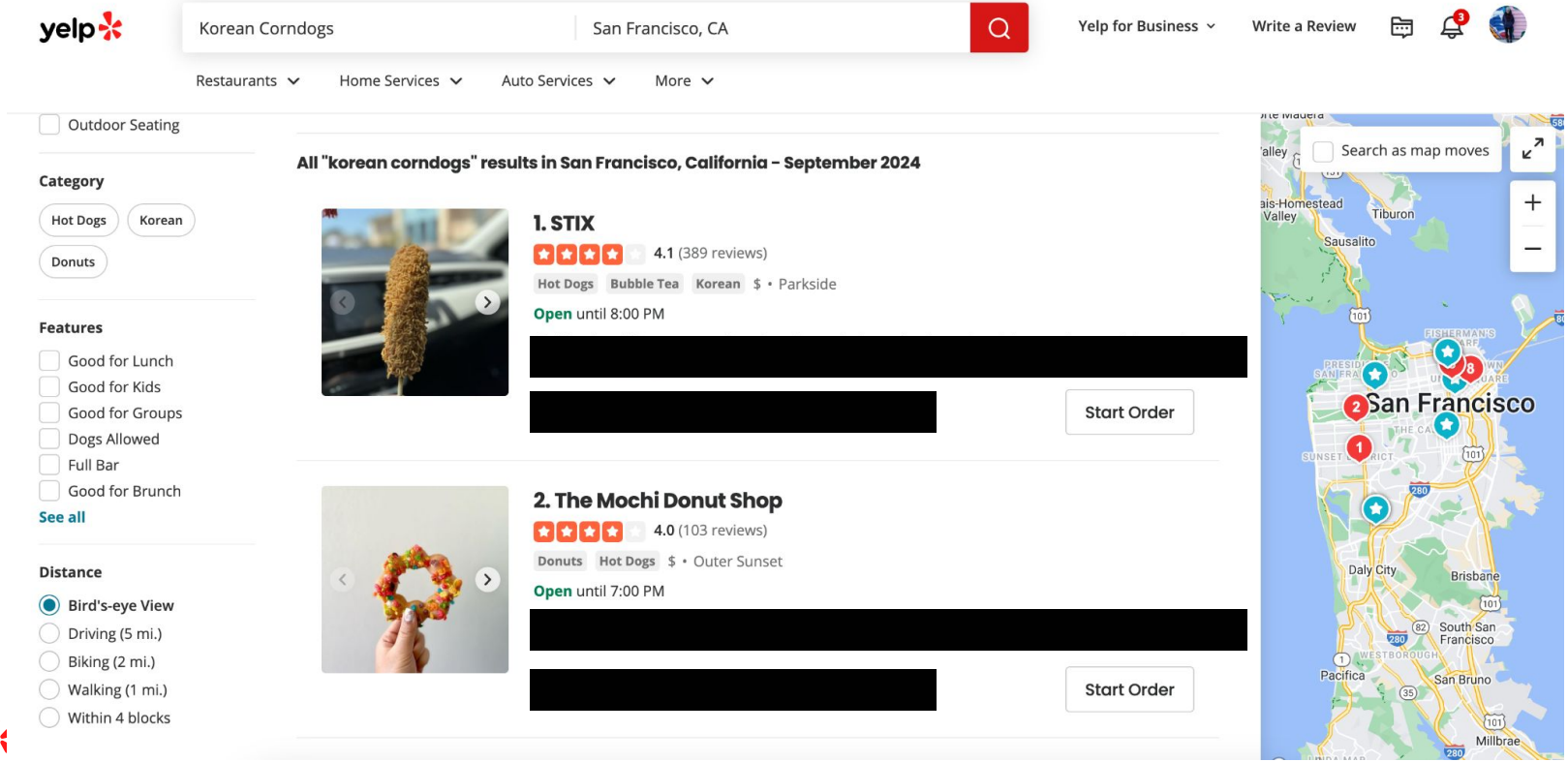
Restaurants Search

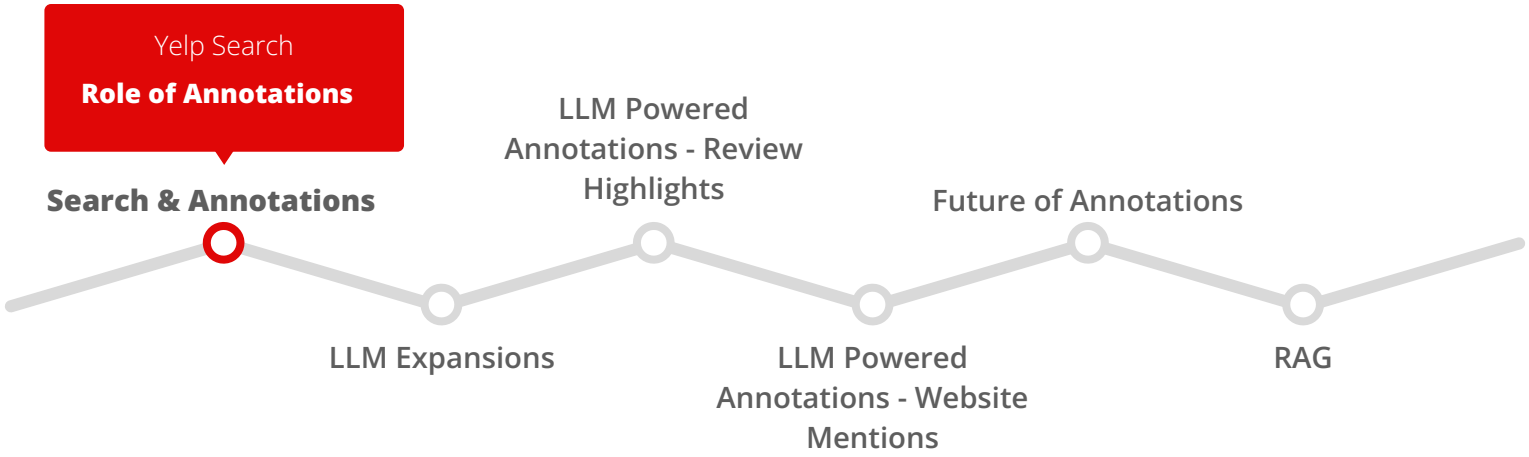


Services Search



Search Engine Results Page (SERP)

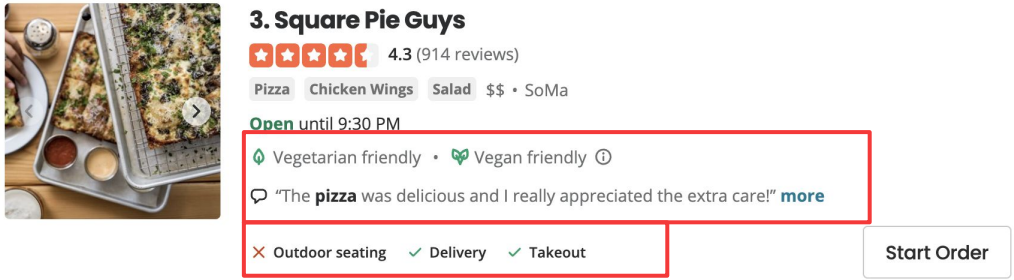


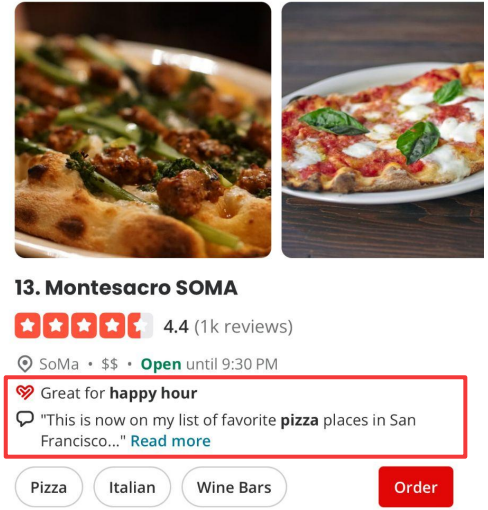

Yelp Search
Role of Annotations
Search & Annotations
LLM Expansions
LLM Powered Annotations - Review Highlights
LLM Powered Annotations - Website Mentions
Future of Annotations
RAG

Annotations

Additional information about a search result

- Relevance proof
- Unique features / differentiators
- Personalized appeal





Review Highlights

- A type of annotation
- Spans of review text, often with a bolded word or phrase

Nov 13, 2023

Amazing pizza and pasta place in the sunset district!

The atmosphere for the restaurant was great! It is perfect for a date spot and even for a casual meal. They have options to sit outside as well inside. However the wait time could definitely get long.

I got the Margarita Pizza, Vodka Pasta, and Cacio de Pepe Pasta. The pasta was pretty normal and nothing too amazing. The cream for each was extremely evident and was pretty heavy. As a result I wouldn't really recommend the pastas.

The pizza was a different story. The pizza was truly remarkable and outstanding. In the Bay Area there are few to many restaurants that have good pizza. It was light and extremely flavorful. Although it was a simple pizza, the sauce and the way the cheese melted really allows the pizza to stand out. The dough was cooked very well and was a perfect mix of crunchy and chewy.

Overall it was a delight! I would definitely recommend it to anyone, as I will definitely be going back!

25. Fiorella - Sunset

4.2 (363 reviews)

Italian · Pizza · Cocktail Bars · $$$ · Inner Sunset

Closed until 5:00 PM

"The **pizza** was truly remarkable and outstanding." **more**

✓ Outdoor seating ✓ Delivery ✓ Takeout

What makes a good annotation?

Yelp has a dual responsibility

- Portray the business accurately
- Be clear to the user, surfacing useful information

Prioritization of Annotation types

01 Confirmation of business relevance

02 Something interesting or unique about the business

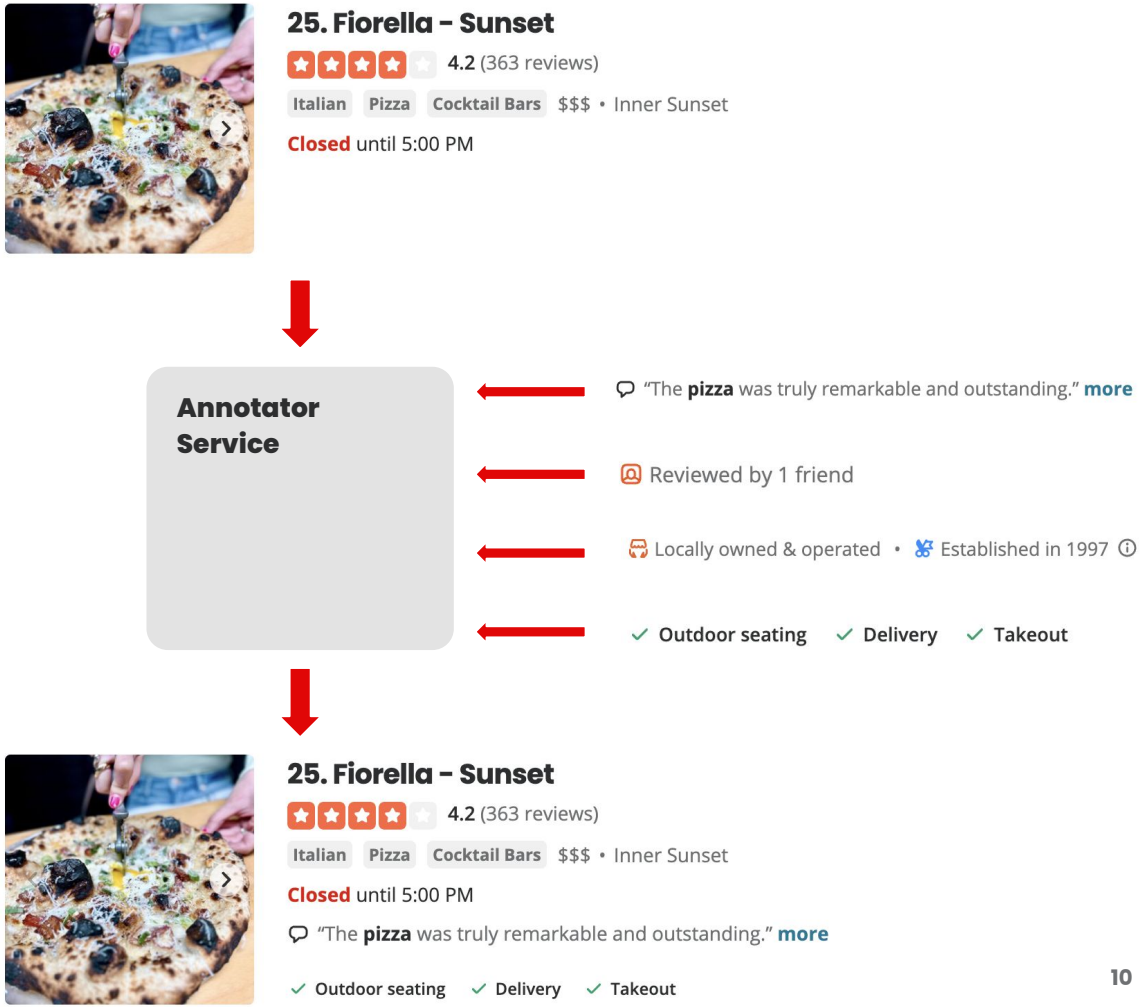
03 Supplementary information about the business

Annotator Service

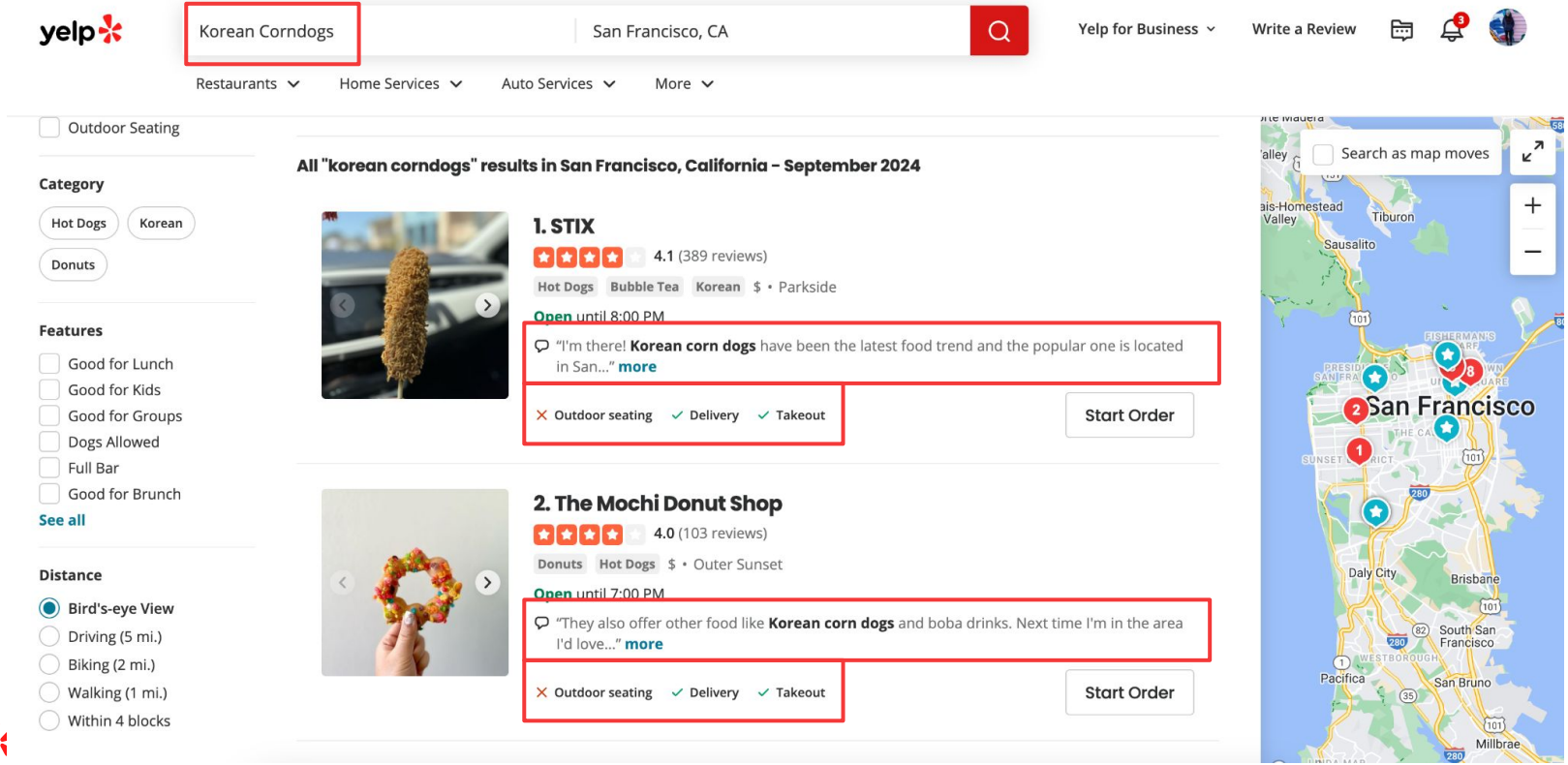
Given a business, search query, and Yelp user

Determine the best annotation(s) to show for the result

Based on the rough principles laid out previously



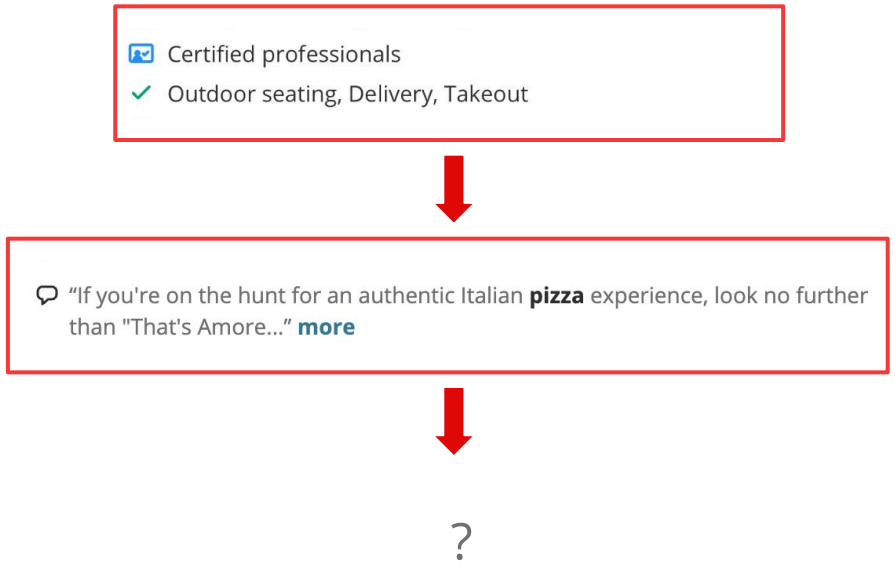
Relevance Proof

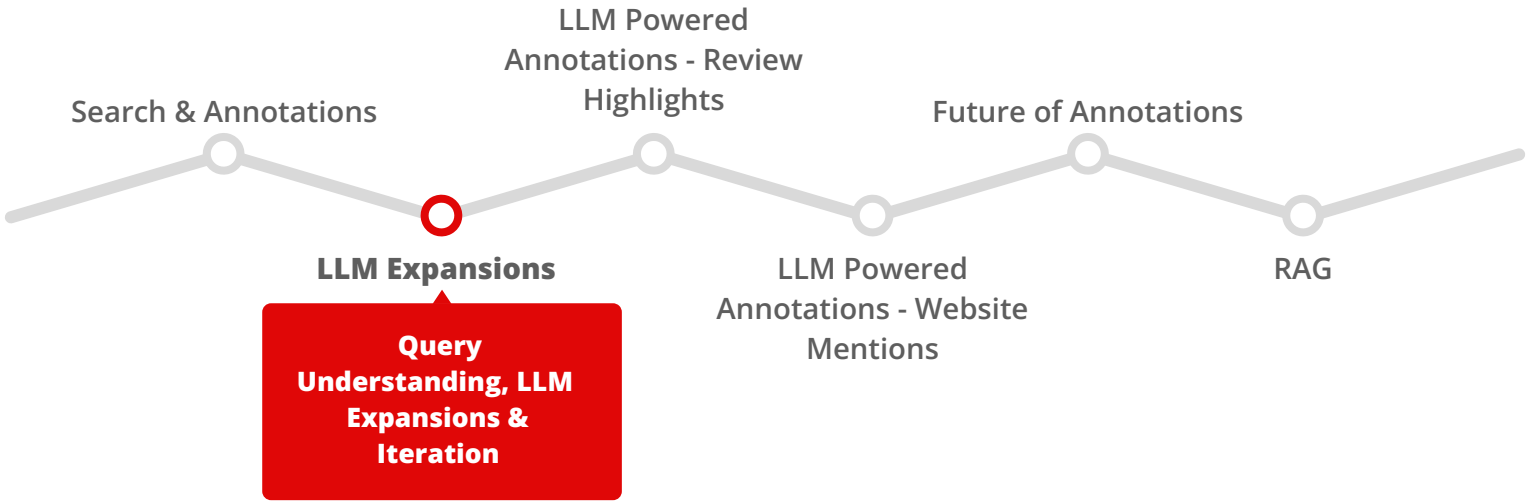


Annotation Roadmap

Key considerations as we make progress:

- Leverage community content
- Right balance between LLMs, manual work, simple coded strategies
- Keep iterating to improve intelligence of annotations




Search & Annotations
LLM Powered Annotations - Review Highlights
Future of Annotations
LLM Expansions
LLM Powered Annotations - Website Mentions
RAG
Query Understanding, LLM Expansions & Iteration

Query Understanding

First step in developing LLM expansions is understanding user intent.

What is the user looking for

- A specific category of businesses
- Particular dish or service
- A specific business or chain
- Broad query including many types of businesses
- etc

"Dentists"

"Mango sticky rice"

"Custom landscaping"

"McDonalds"

"Things to do as a tourist in London"

"restaurant for a wedding reception"

"Movers that can handle grand pianos"

LLM Query Expansions

Given a search query, which phrases are most important to highlight from business reviews?

Key Phrases:

- Search query -> Starbucks
- Categories -> Cafes, Breakfast
- Offerings -> Coffee, tea, muffins

Key Phrases:

- Search query -> affordable hair salon
- Adjectives -> affordable, cheap, value for money
- Categories -> hair salon
- Offerings -> haircut, blowout, styling, hair dyeing

Query Expansion Iterations

Multiple iterations of this process

Modularity - expansions can be plugged into various systems

Iteration and scaling discussed more in a Haystack US talk (QR code below)

- May 2022

  Query: healthy food

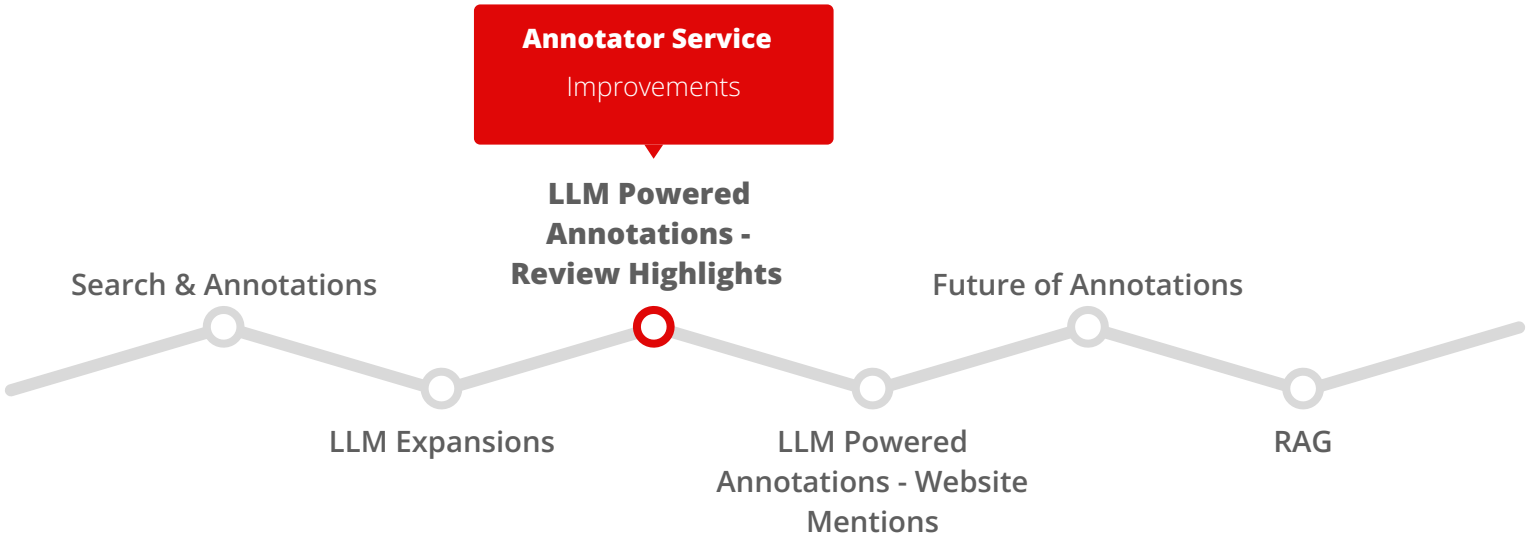
  Key concepts: healthy food, healthy, organic

- March 2023

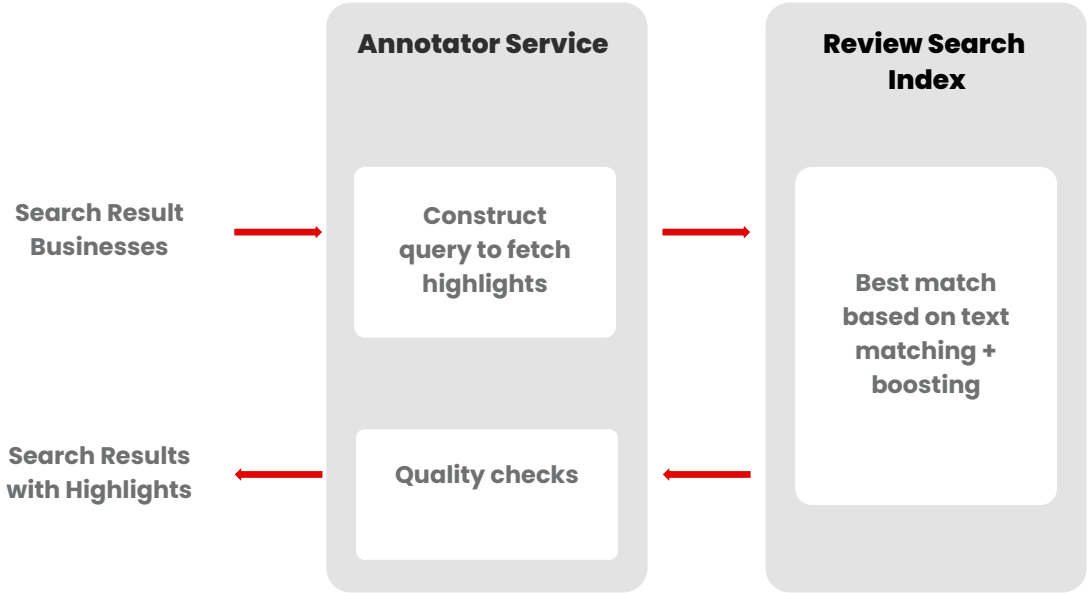
  healthy food -> healthy food, healthy, organic, low calorie, low carb

- September 2023

  healthy food -> healthy food, healthy options, healthy | nutritious, organic, low calorie, low carb, low fat, high fiber | fresh, plant-based, superfood


Annotator Service
Improvements
LLM Powered Annotations - Review Highlights
Search & Annotations
LLM Expansions
LLM Powered Annotations - Website Mentions
Future of Annotations
RAG

Review Highlights Infrastructure



Areas of Improvements

Query Understanding

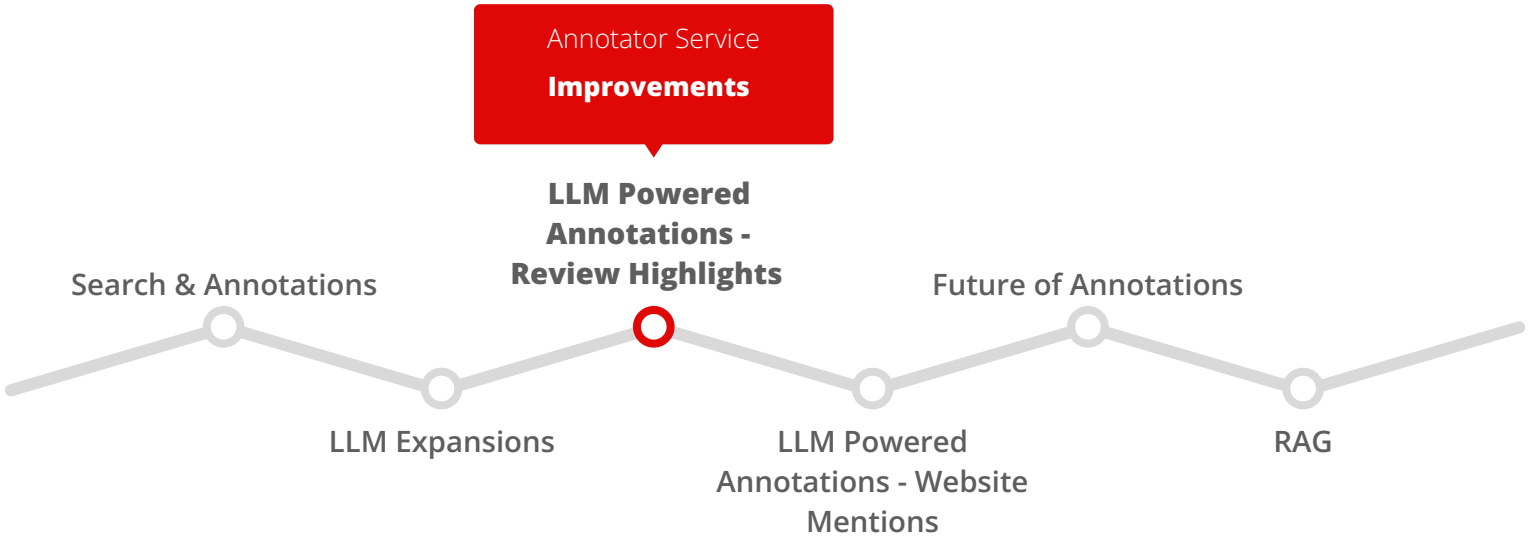
- Matching on search query only
- No strategy for unseen queries

Visibility into Selection

- Single top match doesn't allow any extra control

Handling Text Data

- Stemmer reacts unexpectedly to some phrases
- Highlights clipped incoherently


Annotator Service
Improvements
LLM Powered Annotations - Review Highlights
Search & Annotations
LLM Expansions
LLM Powered Annotations - Website Mentions
Future of Annotations
RAG

Improvements

Shortcomings	Improvements
Matching on search query only	Recall based on search query and relevant phrases
No strategy for unseen queries	Heuristic for unseen queries

Query Understanding

LLM Expansions in Highlights

Integrated LLM expansions by recalling on relevant phrases instead of just search queries

Heuristic for unseen queries

~8x increase in number of phrases to match on per search

```
healthy food -> healthy food, healthy options, healthy
| nutritious, organic, low calorie, low carb, low fat,
high fiber | fresh, plant-based, superfood
```

"Great place for **healthy food** options. They have pre-made salads and sandwiches." **more**

"Packed full of **healthy** vegan goodness, the Buddha bowl of rice and veggies feels both virtuous and..." **more**

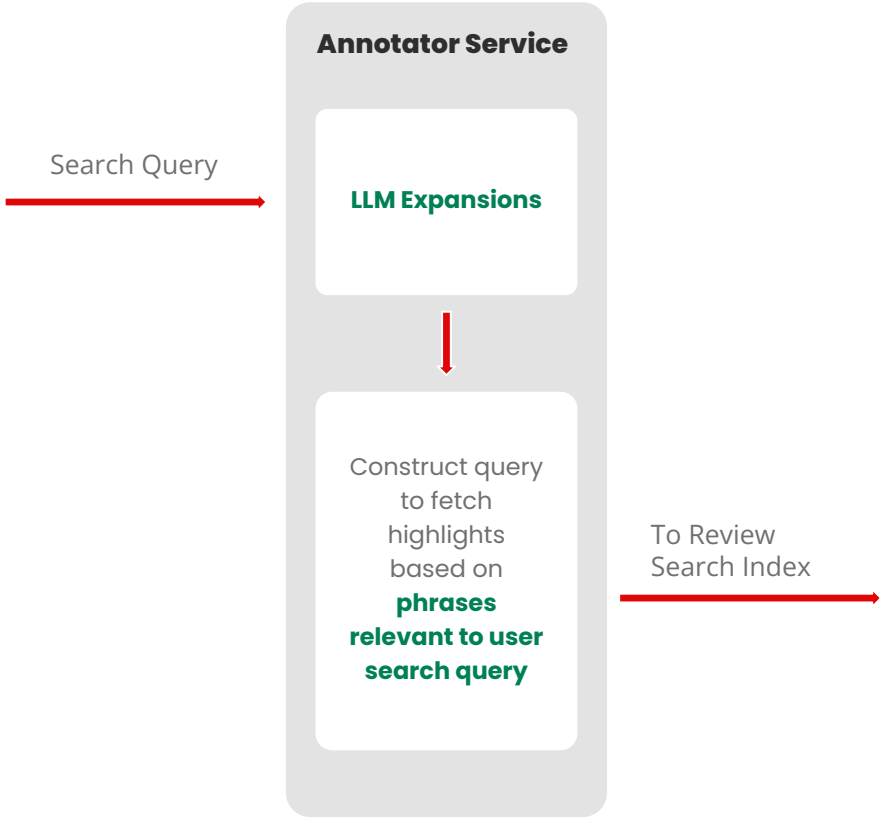
"I never knew vegan food can be so **nutritious** and delicious." **more**

"Falafel were also among the best I ever had, **fresh**, moist, perfect. Roast aubergine also excellent..." **more**

Query Understanding

LLM Expansions in Highlights

Easy to slot into our existing architecture without an overhaul



Visibility into Selection

Improvements

Shortcomings

Single top match doesn't allow any extra control

Improvements

Top 5 best matches returned, allowing smart ranking

Filtering based on quality checks

Visibility into Selection

Highlight Scoring Model

Intuitively, we care about - recency, relevance of phrase, grammatical correctness, while trying to represent the business rating accurately

Smart ranking model uses those features, trained on hand labeled data.

"Korean corn dogs" -> "Korean corn dogs", "corn dogs"

Examples:

[2020] I did an online order to get my family some **Korean corndog**

vs

[2023] A great place in SF to get your **Korean corndog** fix

[2023] This spot really curbs the craving for authentic **Korean corn dogs**

vs

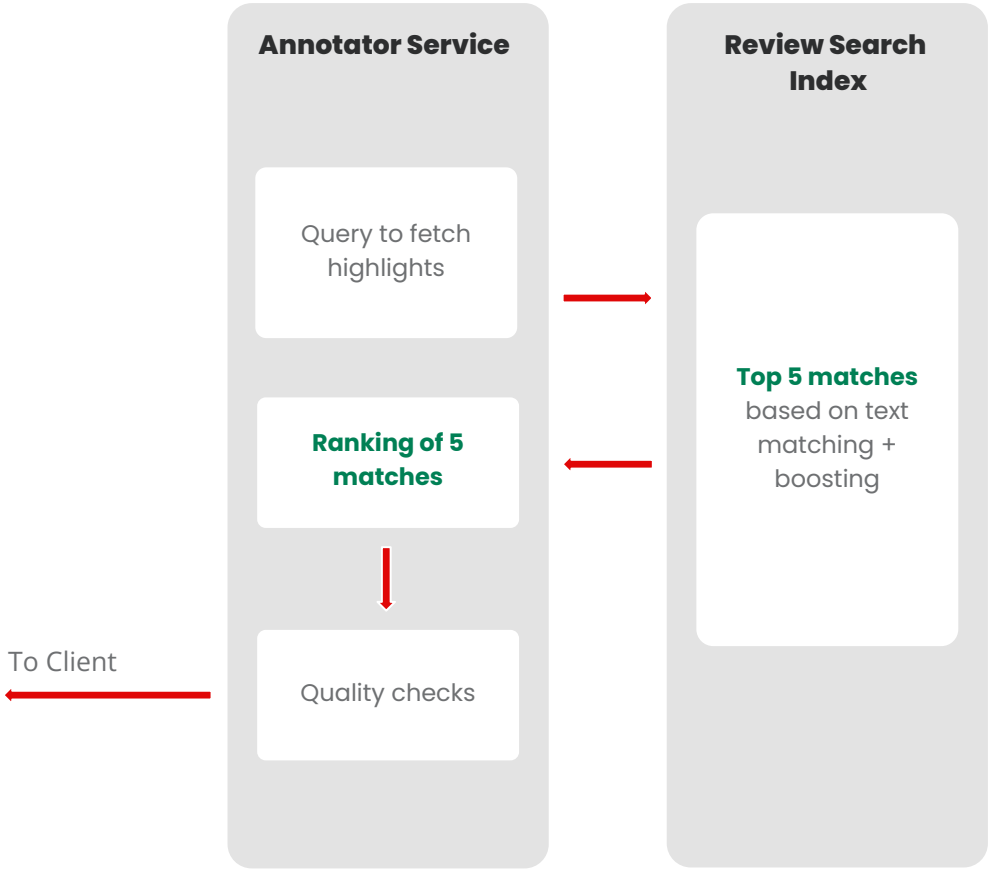
[2023] The **corndog** and wasabi mayo was a delicious pairing

Visibility into Selection

Highlight Scoring Model

Returning multiple matches instead of single match allows us to incorporate smart ranking

Easy to slot into our existing architecture without an overhaul



Filtering Inappropriate Content

We filter out highlights with inappropriate content, or with bad formatting

A GREAT PLACE IN SF to GET your KOREAN CORNDOG fix

Call us at +1 94560362973 for Korean corndogs

Any highlights with NSFW content

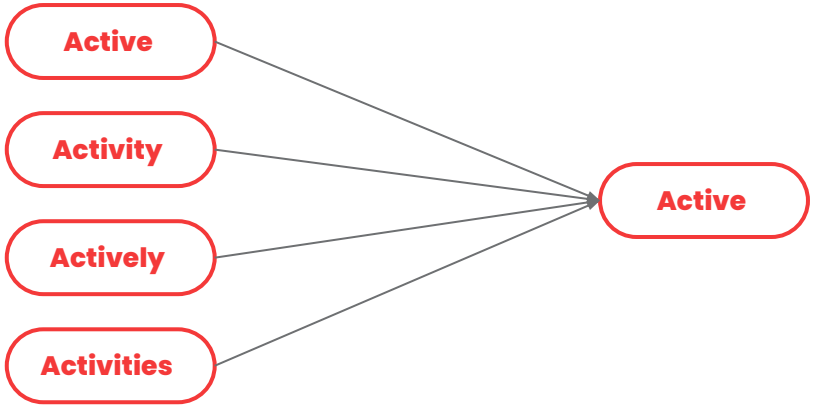
Improvements

Shortcomings	Improvements
Highlights clipped incoherently	Enforce matching on sentence boundaries
Stemmer reacts unexpectedly to some phrases	Special handling of hard to stem phrases

Handling Bad Stemming

Phrases like “active” stemmed to the same root as “activities” resulting in incorrect matching and bolding

Maintain hardcoded list of incorrect stemming, use it for filtering



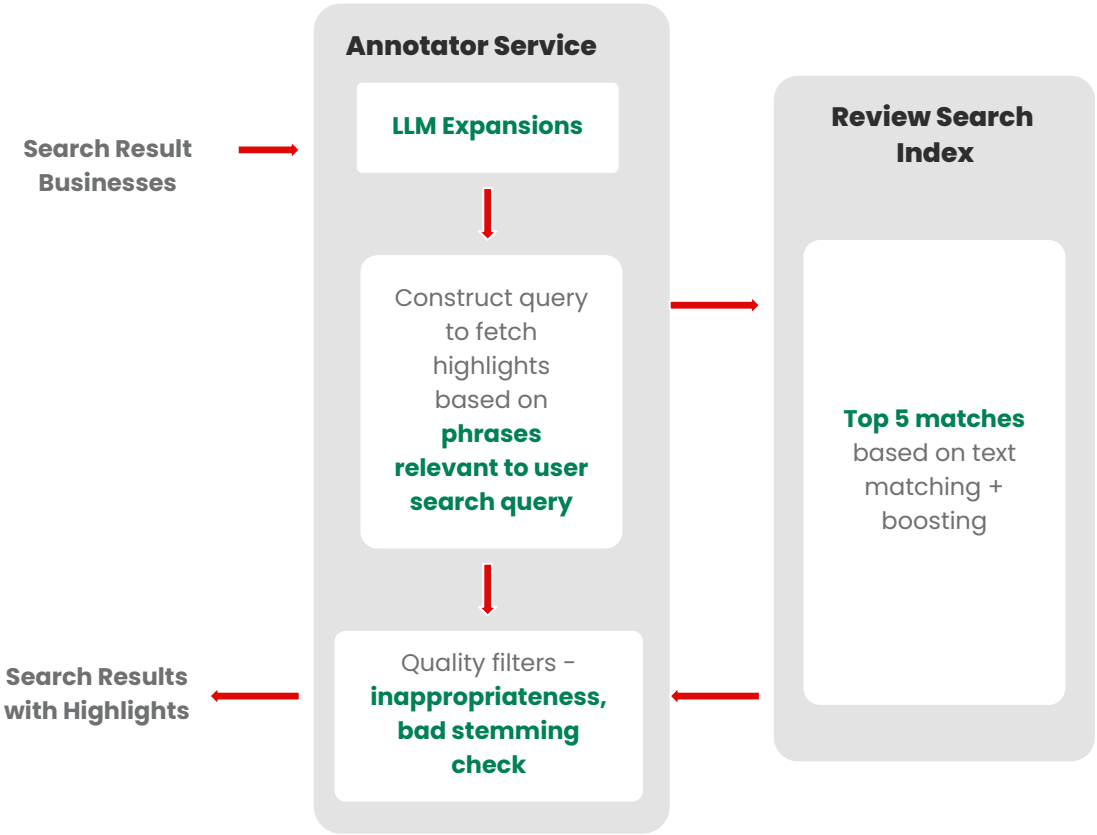
“The plants were healthy, the roads were **active** and the view was beautiful.” **more**

Filter out bad stemming, fall back to a highlight with a less relevant phrase

“It was **fun** to walk around and browse through all the shops: you could spend all the in the area as...” **more**

Result: Smarter Retrieval Flow

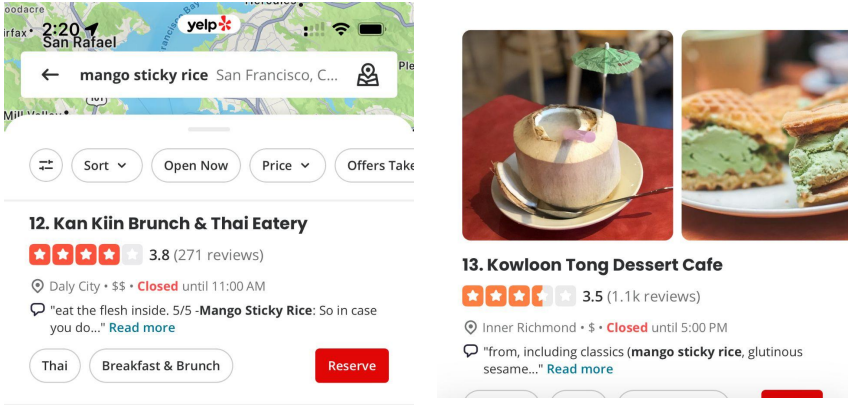
New Highlights Infrastructure



Comparison between old and new

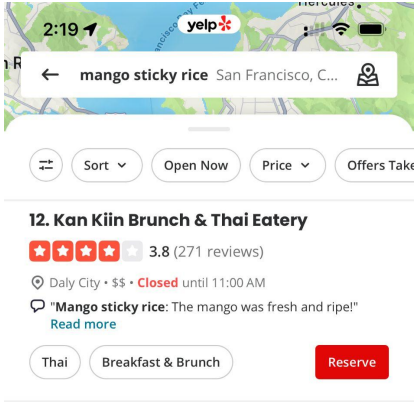
Old System

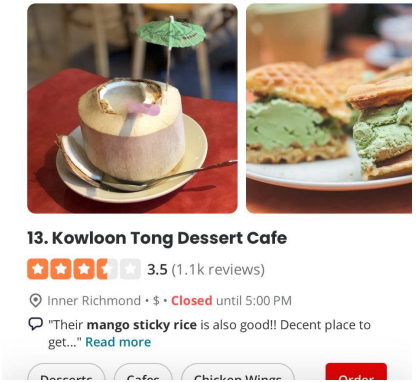
Incomplete sentences, sentiment unclear

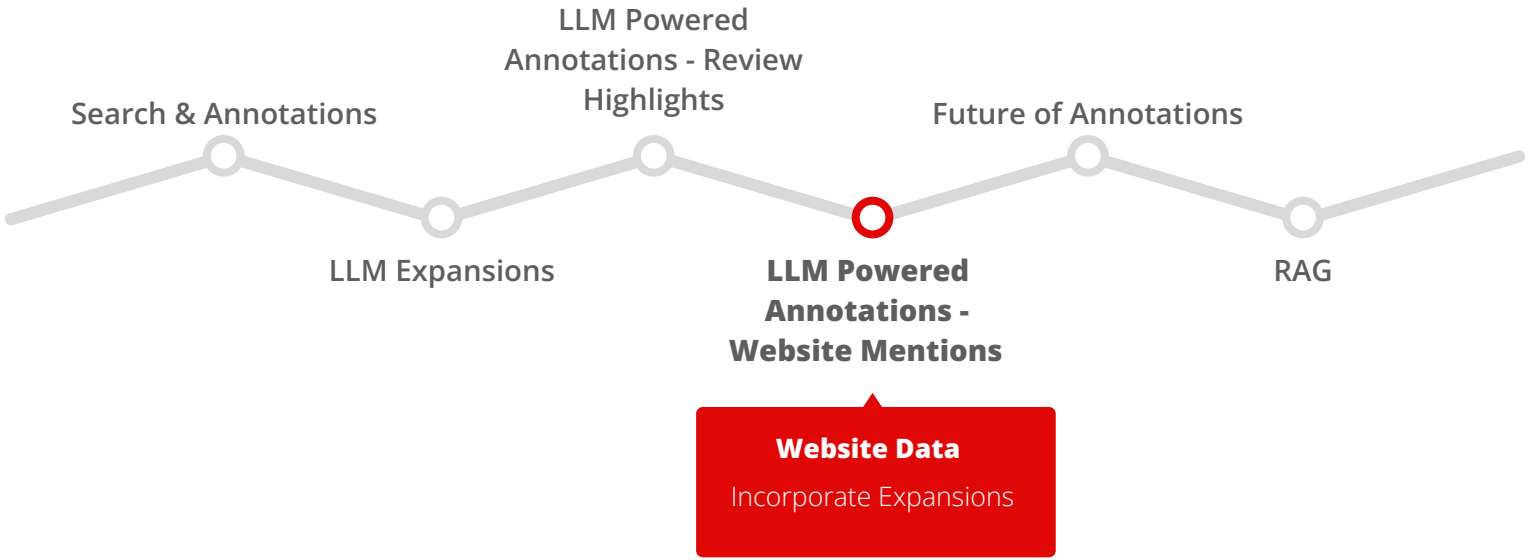


New System

Complete thoughts, sentiment clear





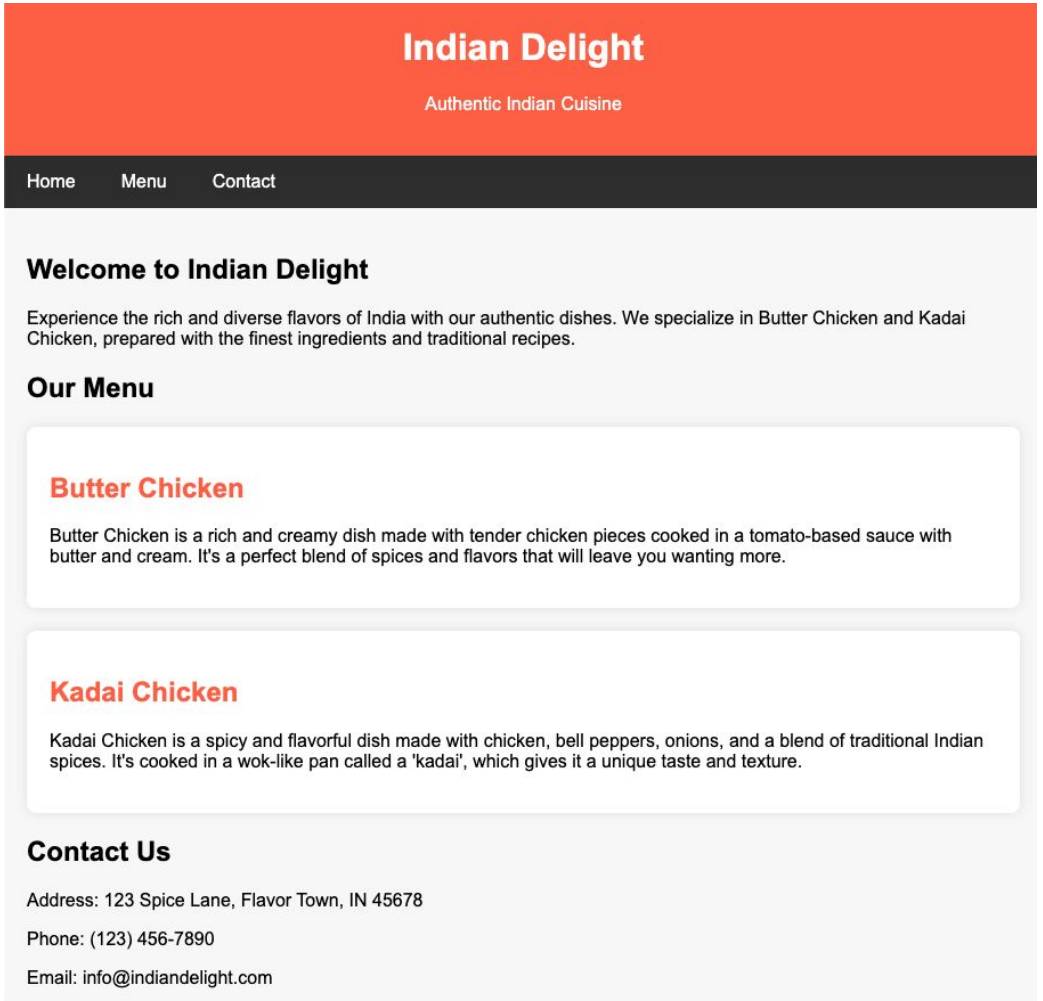

Search & Annotations
LLM Powered Annotations - Review Highlights
Future of Annotations
LLM Expansions
LLM Powered Annotations - Website Mentions
RAG
Website Data
Incorporate Expansions

How can we provide relevance proof for low content or unreviewed businesses?

Website Data

A business with 0 reviews
Website mentions “butter chicken” and “kadai chicken”

Surfacing this to the user provides strong proof of the business’ relevance

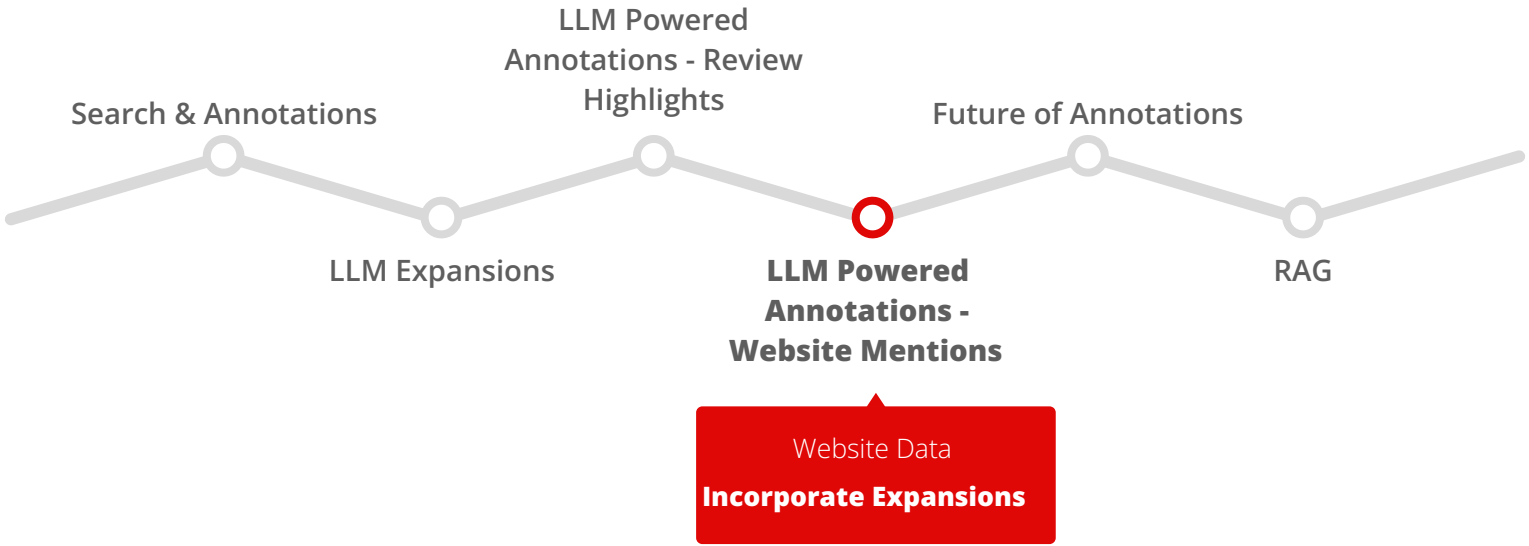


Challenges

- Scale of data
- Information synthesis
- Quality

<h2>Welcome to Crystal Harbor</h2>

Nestled within the vibrant metropolis of Grandoria, Crystal Harbor is a hidden gem that offers an unparalleled blend of urban sophistication and natural splendor. This enchanting destination, spanning over 500 acres, is a sanctuary where the hustle and bustle of city life gives way to serene landscapes, cultural landmarks, and unforgettable experiences.


Search & Annotations
LLM Powered Annotations - Review Highlights
Future of Annotations
LLM Expansions
LLM Powered Annotations - Website Mentions
RAG
Website Data
Incorporate Expansions

Solution

- Use LLM expansions
- Resource of phrases to match on, consisting of all possible search queries and their expansions

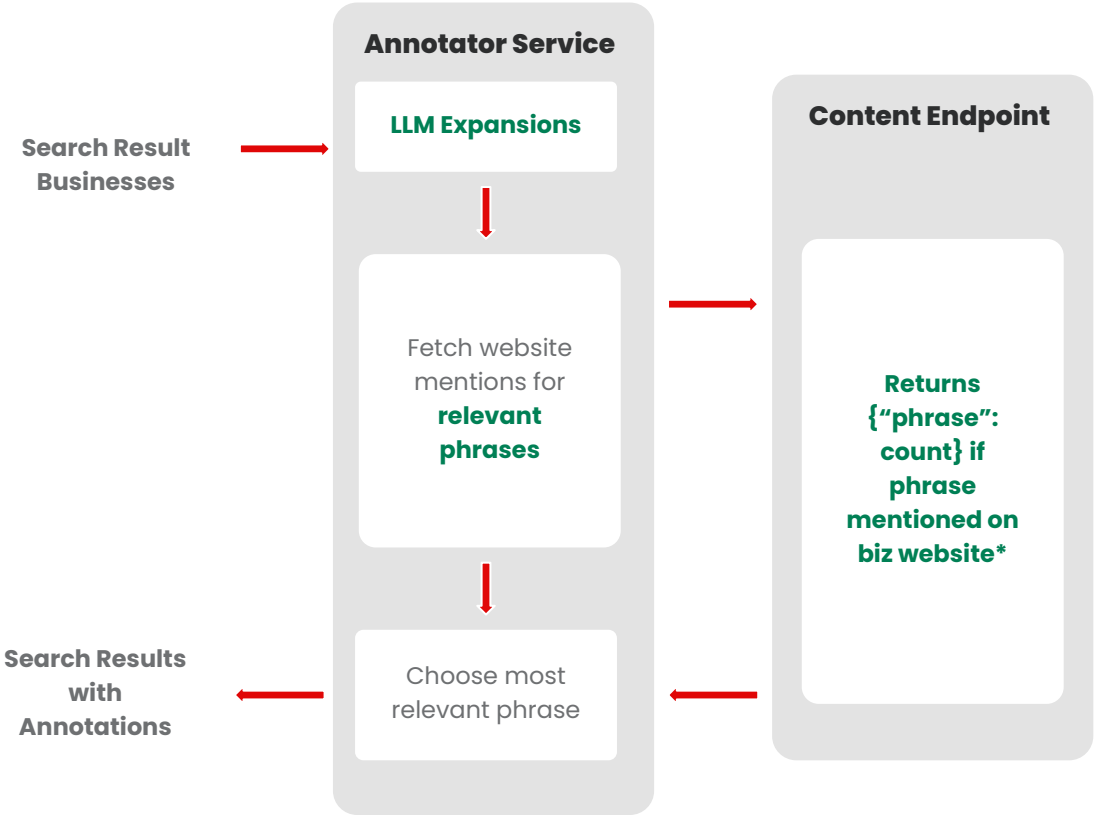
Welcome to Crystal Harbor

Nestled within the **vibrant** metropolis of Grandoria, **Crystal Harbor** is a **hidden gem** that offers an unparalleled blend of **urban** sophistication and **natural** splendor. This enchanting **natural destination**, spanning over 500 acres, is a **sanctuary** where the hustle and bustle of city life gives way to serene **landscapes**, **cultural landmarks**, and **unforgettable experiences**.

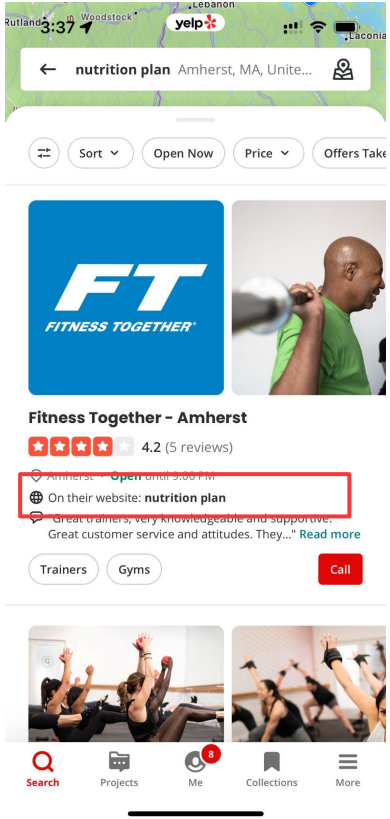
{**vibrant: 1**, **Crystal Harbor: 2**, **hidden gem: 1**, **urban: 1**, **natural: 2**, **destination: 1**, **sanctuary: 1**, **landscapes: 1**, **cultural landmarks: 1**, **unforgettable experiences: 1**}

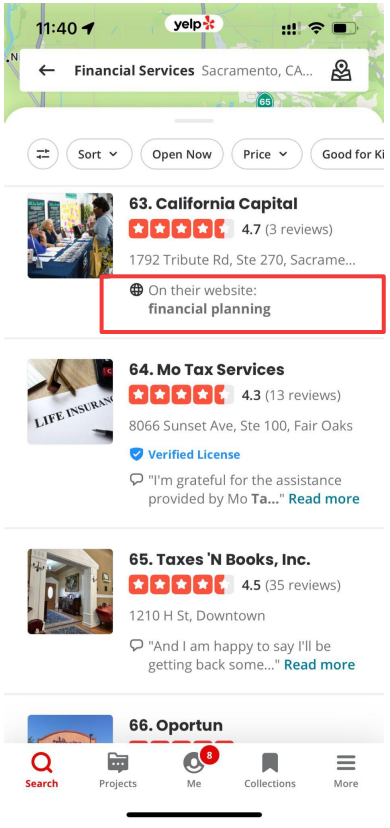
Website Mentions Infrastructure

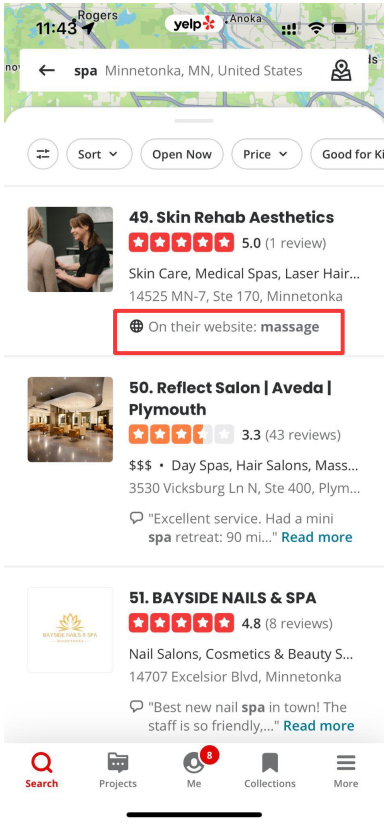
* Only for businesses that provide Yelp with their website URL

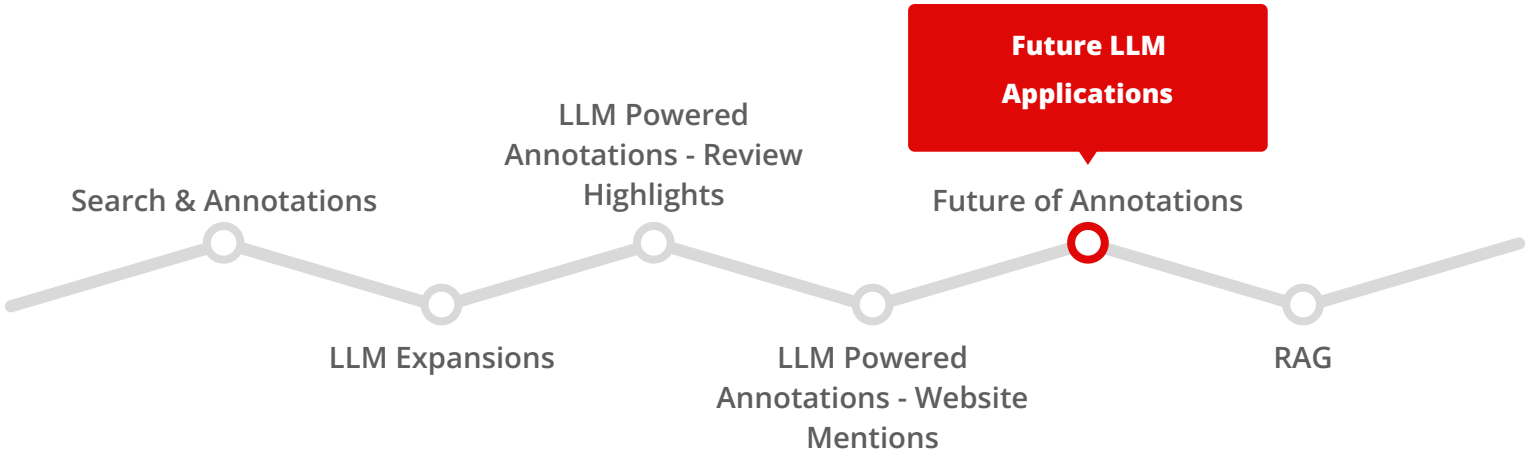


Examples


nutrition plan Amherst, MA, Unite...
Sort
Open Now
Price
Offers Take
FITNESS TOGETHER
Fitness Together - Amherst
4.2 (5 reviews)
On their website: nutrition plan
Great customer service and attitudes. They..." Read more
Trainers
Gyms
Call
Search
Projects
Me
Collections
More


Financial Services Sacramento, CA...
Sort
Open Now
Price
63. California Capital
4.7 (3 reviews)
1792 Tribute Rd, Ste 270, Sacrame...
On their website: financial planning
64. Mo Tax Services
4.3 (13 reviews)
8066 Sunset Ave, Ste 100, Fair Oaks
Verified License
"I'm grateful for the assistance provided by Mo Ta..." Read more
65. Taxes 'N Books, Inc.
4.5 (35 reviews)
1210 H St, Downtown
"And I am happy to say I'll be getting back some..." Read more
66. Oportun
Search
Projects
Me
Collections
More


spa Minnetonka, MN, United States
Sort
Open Now
Price
49. Skin Rehab Aesthetics
5.0 (1 review)
Skin Care, Medical Spas, Laser Hair...
14525 MN-7, Ste 170, Minnetonka
On their website: massage
50. Reflect Salon | Aveda | Plymouth
3.3 (43 reviews)
$$$ • Day Spas, Hair Salons, Mass...
3530 Vicksburg Ln N, Ste 400, Plym...
"Excellent service. Had a mini spa retreat: 90 mi..." Read more
51. BAYSIDE NAILS & SPA
4.8 (8 reviews)
Nail Salons, Cosmetics & Beauty S...
14707 Excelsior Blvd, Minnetonka
"Best new nail spa in town! The staff is so friendly,..." Read more
Search
Projects
Me
Collections
More


Search & Annotations
LLM Expansions
LLM Powered Annotations - Review Highlights
LLM Powered Annotations - Website Mentions
Future of Annotations
Future LLM Applications
RAG

LLM Expansions & Phrase counts

Re-usability of approaches and signals

Future explorations:

1. Other applications of phrases
    a. On the menu: <phrase>
    b. X reviews mention <phrase>

Website mentions: {**vibrant: 1**, **Crystal Harbor: 2**, **hidden gem: 1**, **urban: 1**, **natural: 2**, **destination: 1**, **sanctuary: 1**, **landscapes: 1**, **cultural landmarks: 1**, **unforgettable experiences: 1**}

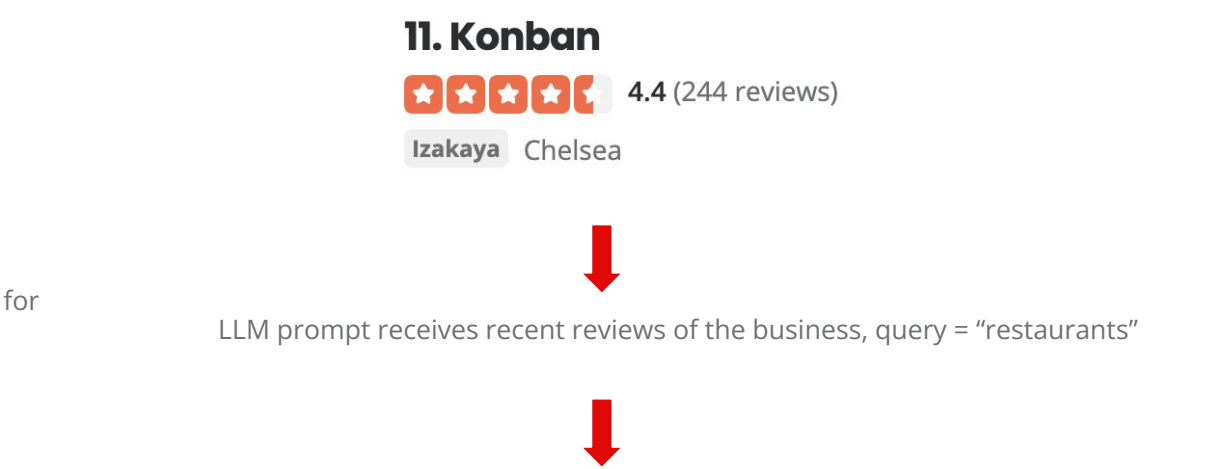
Menu mentions: {**butter chicken: 2**, **kadai chicken: 2**, **chicken curry: 1**}

Review mentions: {**butter chicken: 20**, **kadai chicken: 12**, **chicken curry: 10**, **indian food: 50**, **fresh ingredients: 4**}

LLMs in Highlight Selection

Future explorations:

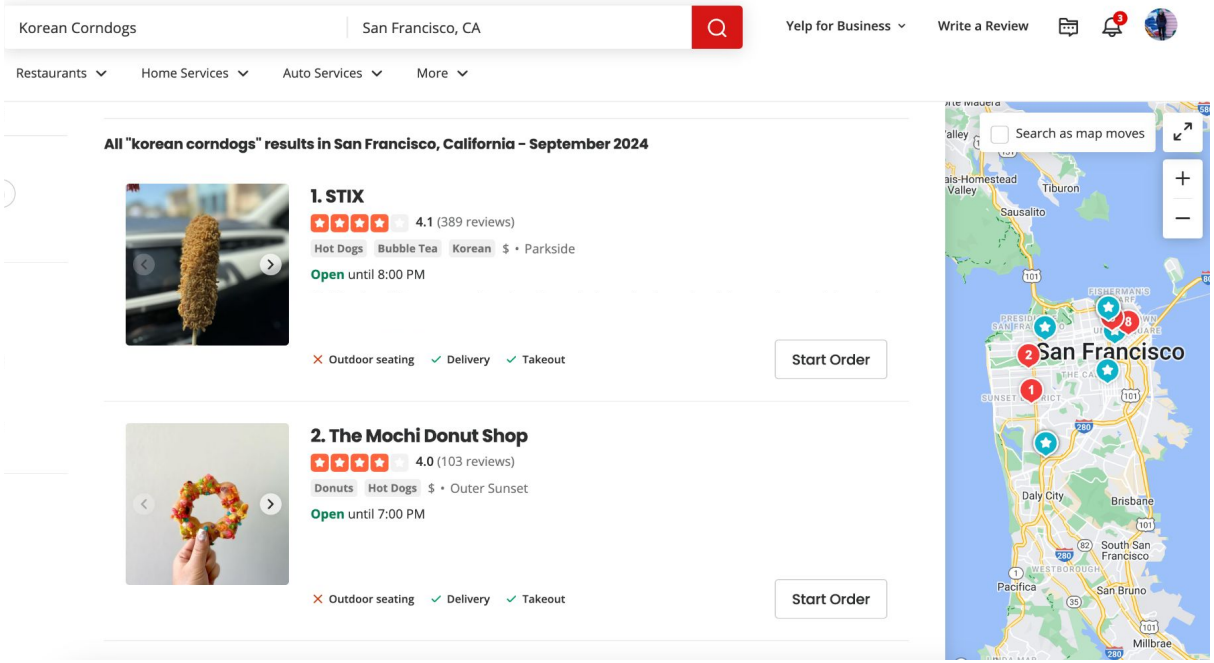
2. Offline precomputation of highlights for biz-query pairs



LLMs in Annotation Selection

Future explorations:

3. Offline precomputation of all annotations for a search page, with awareness of the other results on the page

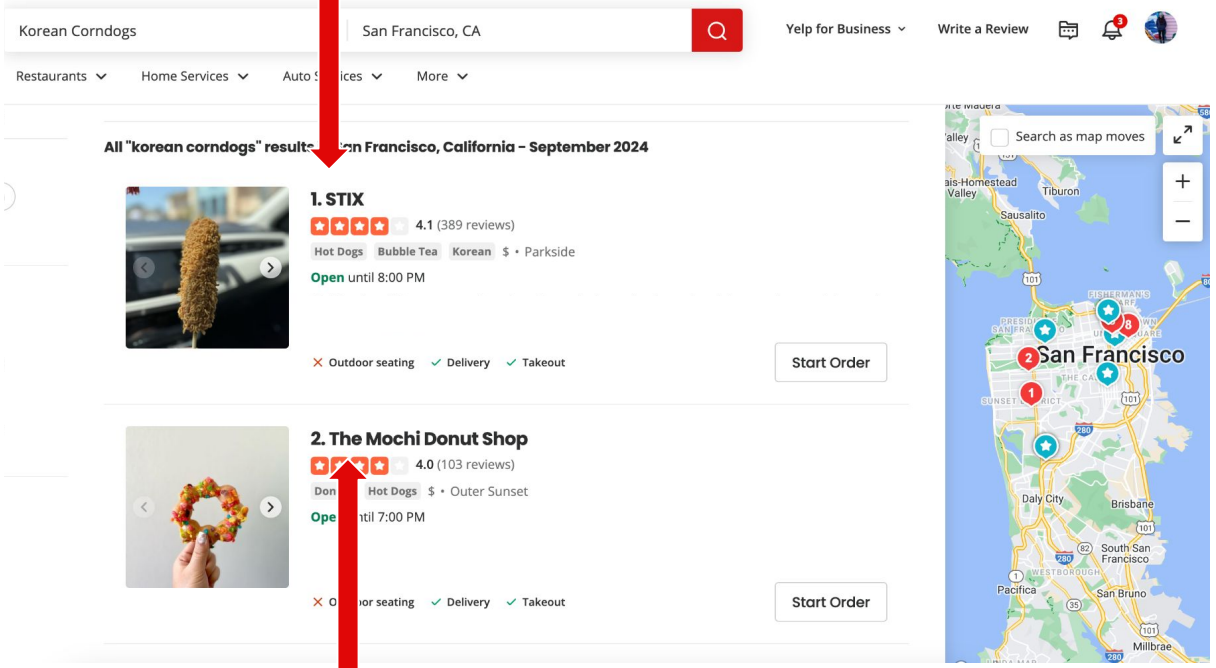


LLMs in Annotation Selection

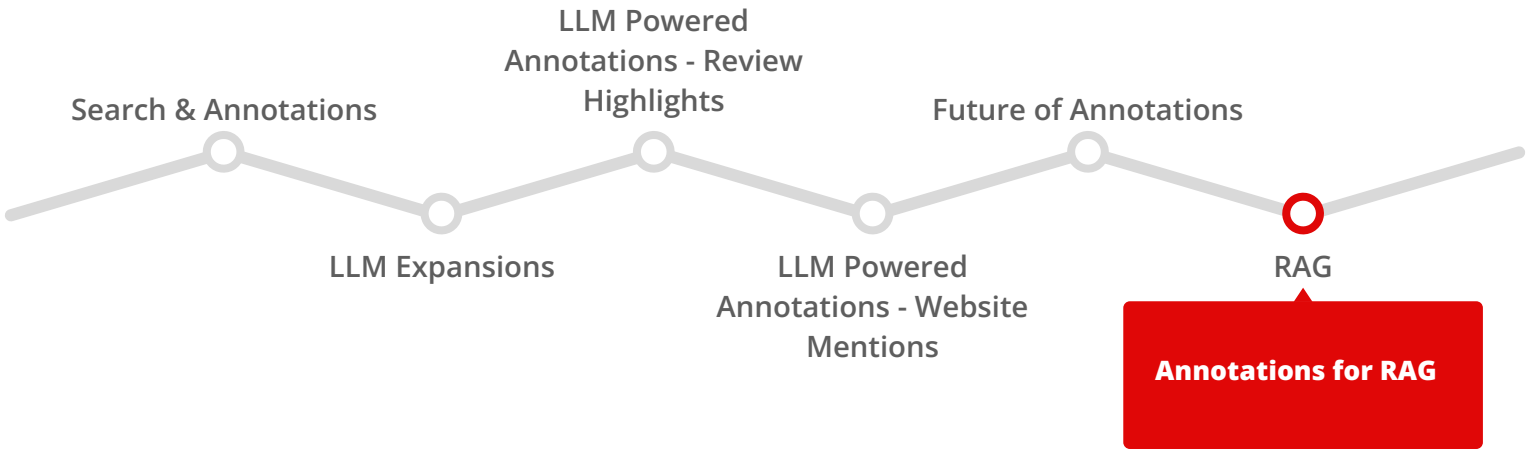
Future explorations:

3. Offline precomputation of all annotations for a search page, with awareness of the other results on the page

Korean corndogs are their **specialty**, users have called them **crunchy** and **satisfying**

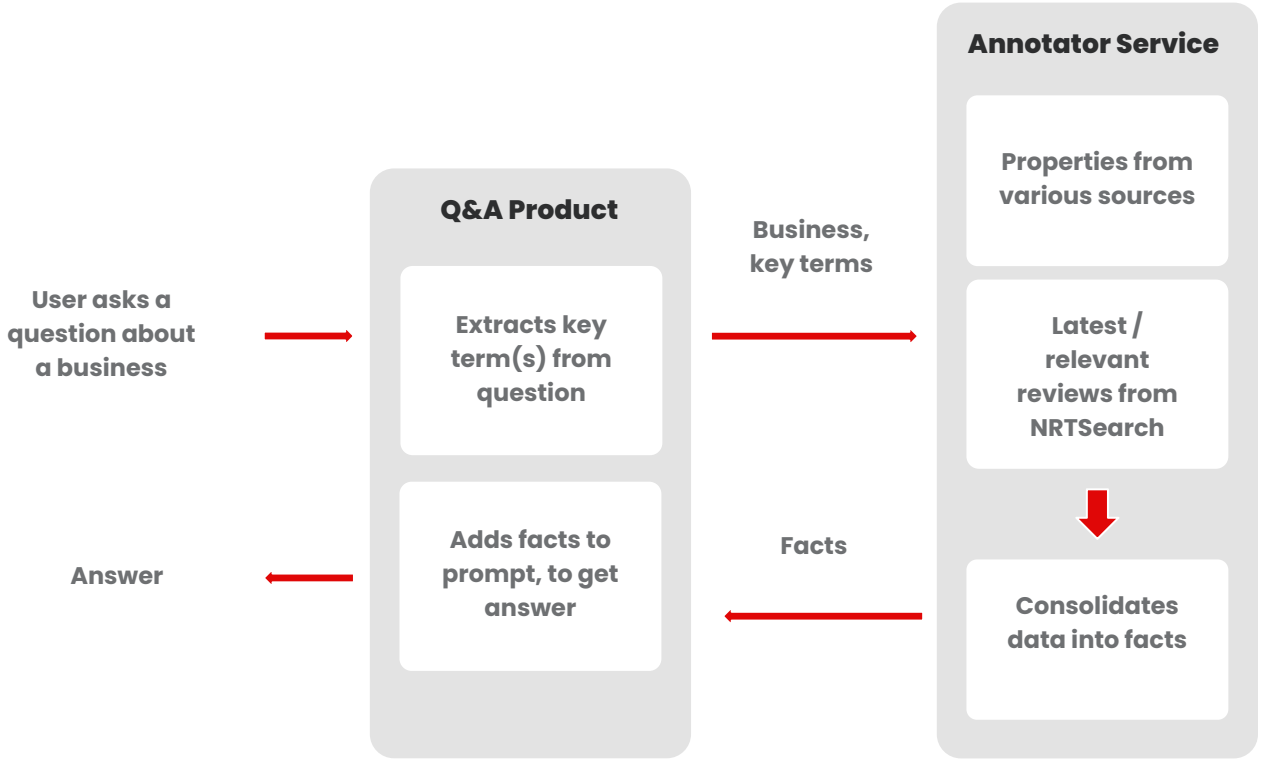


Known more for donuts, but they have **interesting corndog flavors** and beginner friendly **half-and-half options**


Search & Annotations
LLM Expansions
LLM Powered Annotations - Review Highlights
LLM Powered Annotations - Website Mentions
Future of Annotations
RAG
Annotations for RAG

Annotator Service for RAG

- Retrieval-Augmented Generation
- **Powers multiple internal use cases**



THANK YOU!

QUESTIONS?