

Building an Intelligent Assistant

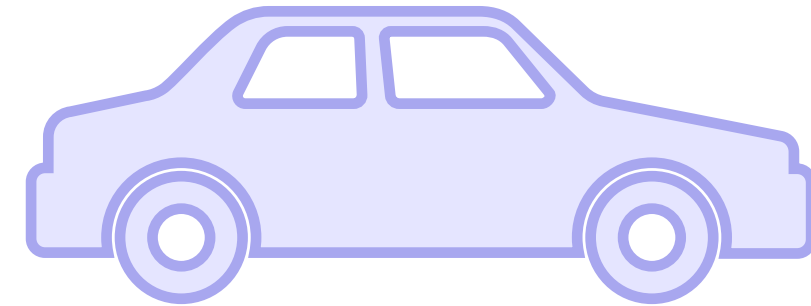
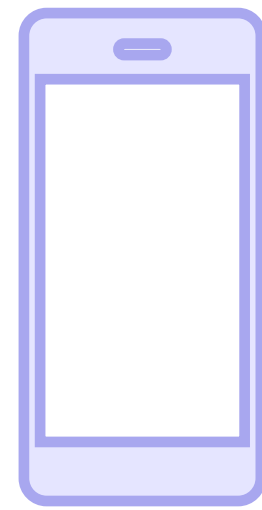
“Tales from the trenches”

Ozlo | [Anant Narayanan](#) | [@anantn](#)

Why?

People are changing their search behavior

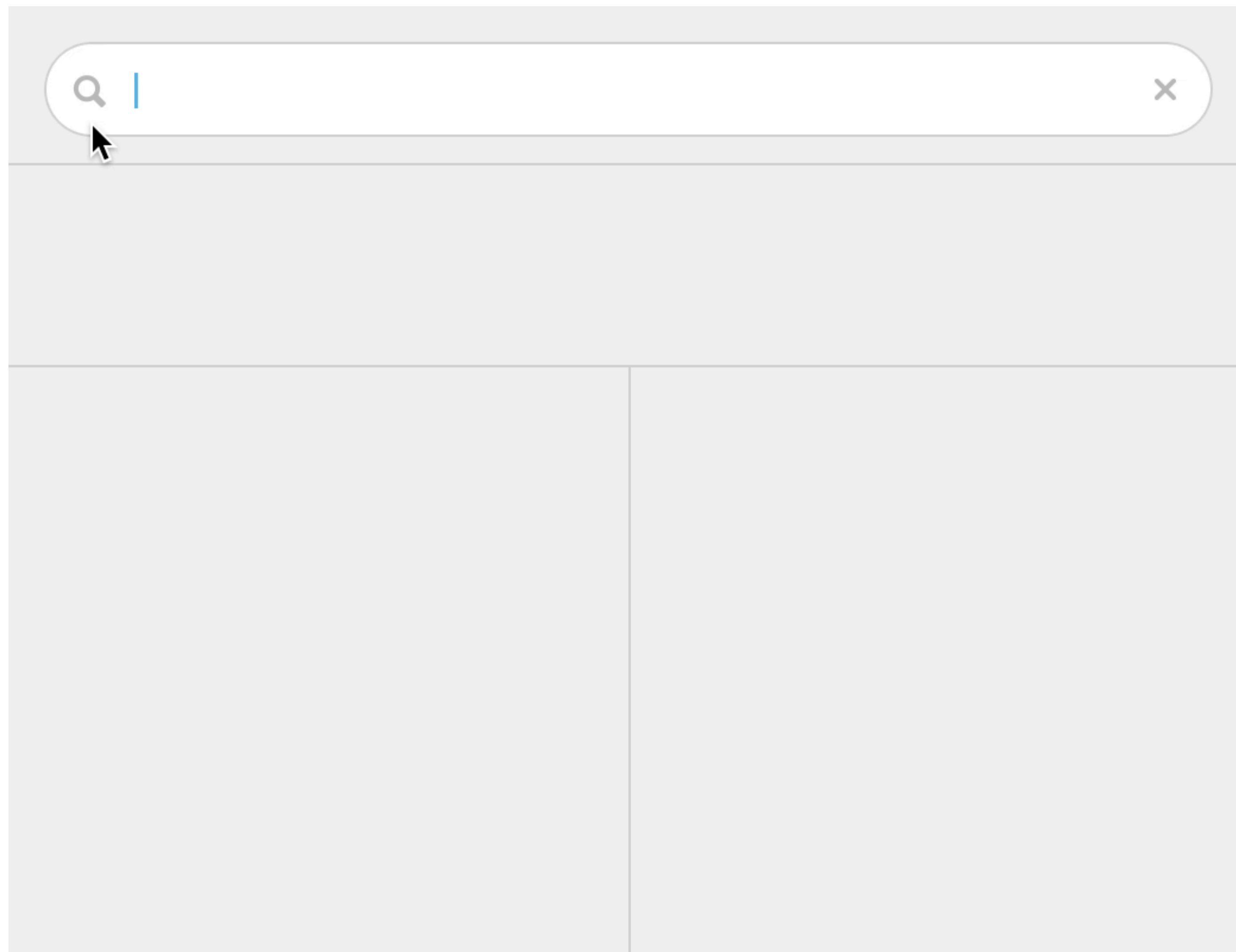
On-the-go



More talking, Less typing



The next frontier is an always available,
conversational, personal,
AI assistant



We've been busy building a
conversational, intelligent
assistant for 3+ years

Here are **10** of the
most important
things we learned

The Building Blocks



KNOW

(facts about the world)



UNDERSTAND

(how people ask questions)



ANSWER

(hold a natural conversation)

Know

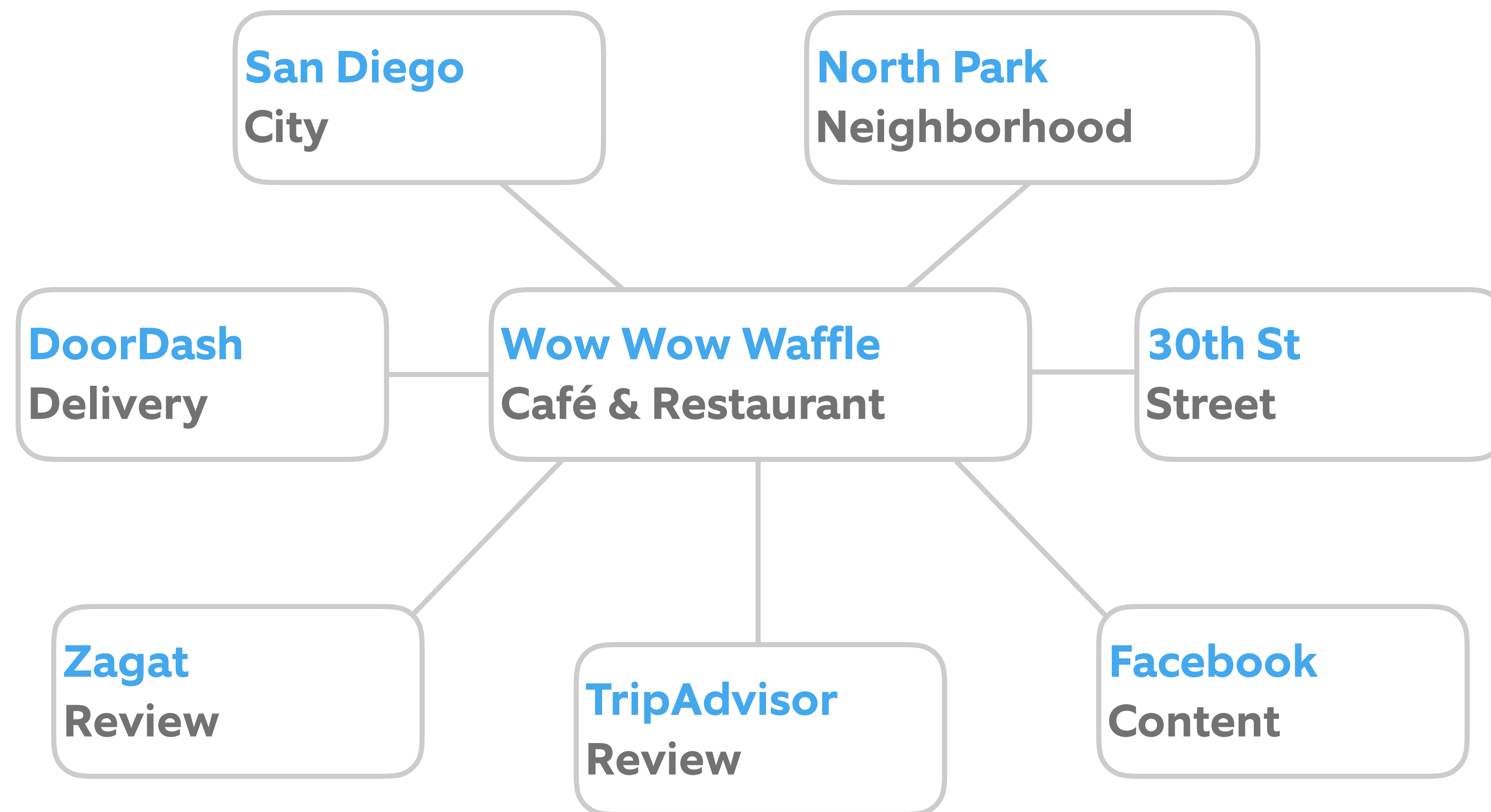


Facts about the World

1

Model the World

The types of things and the relationships between them are finite and countable



2

Build the right kind of data tools

The number of unique entities in the world are also finite and countable

<40M Tracks ever recorded

<5M Famous people in the world

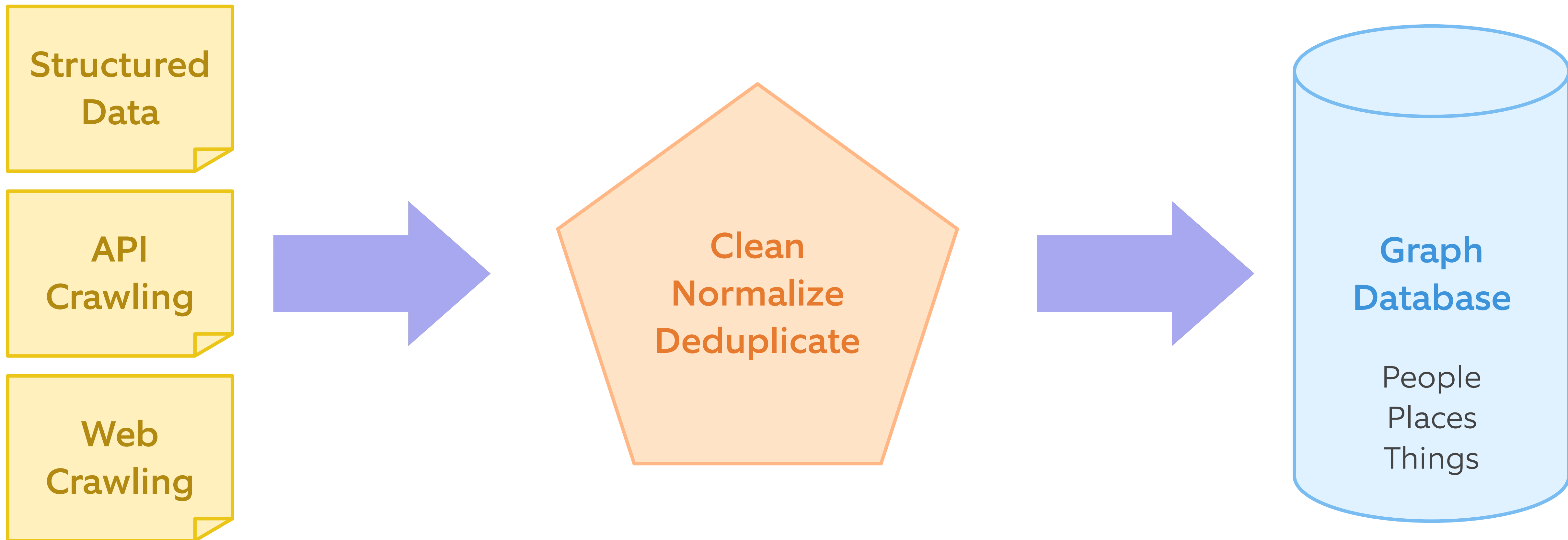
<25M Points of interest in the US

<1M Theatrical movies ever made

<10M Streets in the US

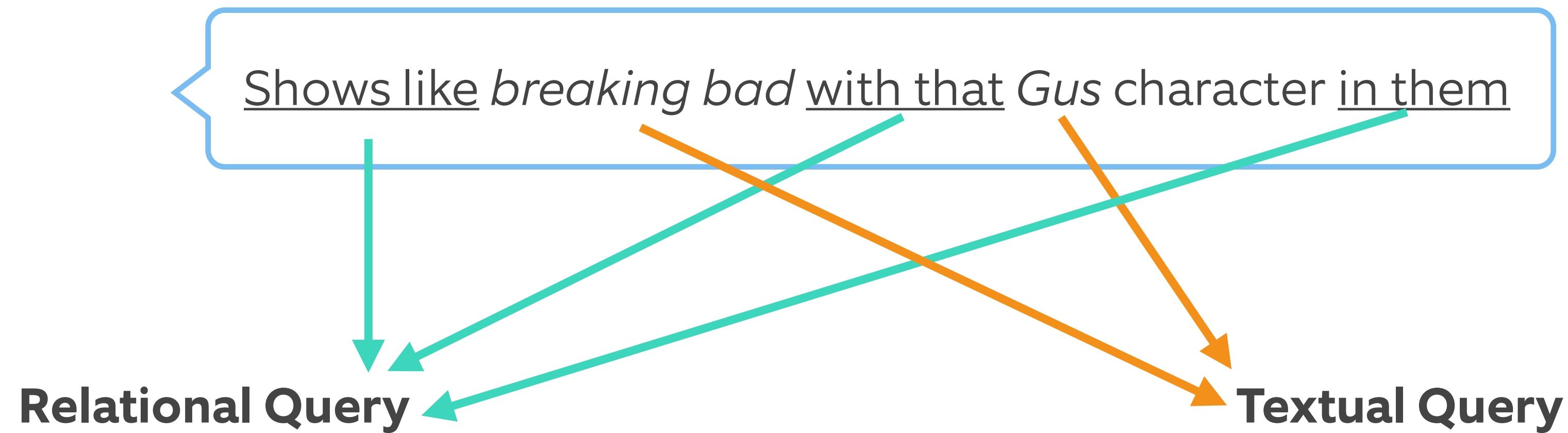
<5K Cities in the world

“Big data” isn’t what it used to be - machines are bigger and faster than ever
Focus on variety of sources (API | Unstructured), speed, ergonomics and data cleanliness



3

Use a hybrid text + graph search engine



Traditional Document ➡ Content IR techniques won't work
Dumping entities into Lucene or ElasticSearch is not the answer

You need a hybrid: text + graph search engine

Understand



How people ask Questions

4

Build a real NLP stack

Recall: Traditional Document ➔ Content IR techniques won't work well

Corollary: Traditional "Noun Phrase" extraction and keyword spotting also won't work well



There are often important hints in the grammar of the utterance. Use them — or risk giving wrong/unsatisfactory answers!

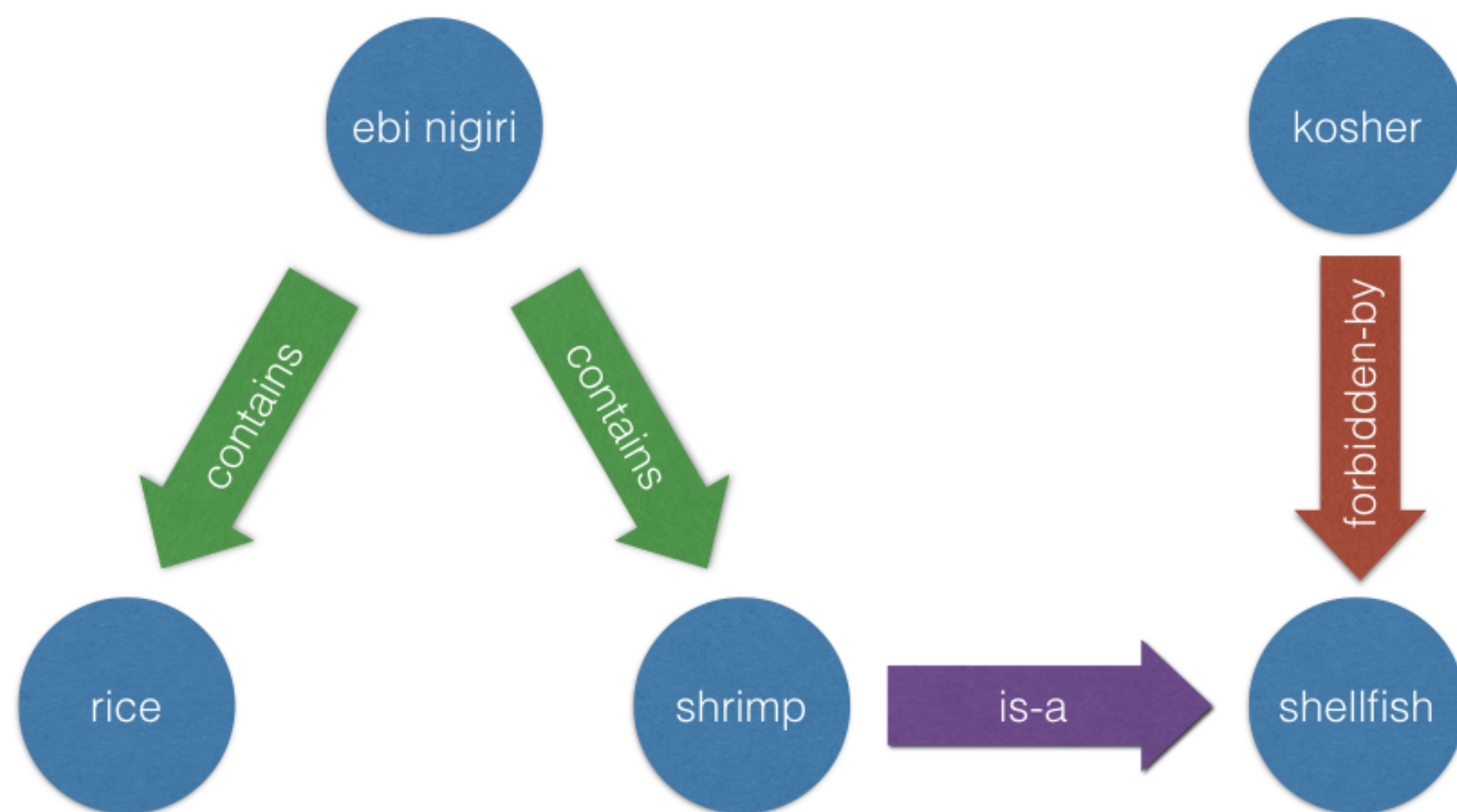
Movies with Tina Fey but without Amy Poehler

Ben Affleck movies that don't suck

5

Integrate “common sense” knowledge

Only looking at the entity graph during query understanding can quickly get limiting:
Make use of a common sense knowledge base



Layer a lexicon or dictionary on top your knowledge to supercharge your query understanding: this is how you get plurals, common variations and synonyms!

Adjective [\[edit \]](#)

kosher (*comparative* **more kosher**, *superlative* **most kosher**)

1. (Judaism) Fit for use or **consumption**, in accordance with **Jewish law** (especially relating to food).
*Only in New York can you find a good, **kosher** hamburger!*
*David's mother kept a **kosher** kitchen, with separate sets of dishes for meat and for dairy.*
*In order for a suit to be **kosher**, it cannot contain both wool and linen together.*
2. (*figuratively, by extension*) In accordance with **standards** or **usual** practice.
*Is what I have done **kosher** with Mr. Smith?*

Antonyms [\[edit \]](#)

- treyf

Related terms [\[edit \]](#)

- kasher
- kashrut

6

Separate *intent* from the *answer*

Which experience feels better?

Chicken soup recipes

Here are a few restaurants near you that serve chicken soup:

Chicken soup recipes

*Sorry, I don't know how to find recipes yet.
Here are a few restaurants near you that deliver chicken soup:*

Build out a very large set of intents, even if you cannot answer them all
Just recognizing them and doing nothing afterwards will improve every other part of your system

Answer



Hold a natural Conversation

7

Embrace Ambiguity

The beauty of a conversation is that system doesn't always have to give an answer immediately

If there is doubt, surface it!

Does Evvia take reservations?

I'm not sure! Here is their phone number:

What are some good coffeeshops on University?

Did you mean University Ave. in Palo Alto or University St. in Menlo Park?

Palo Alto

I found a few good coffeeshops on University Ave. in Palo Alto:

8

Generate natural language responses

Good news: if you have built an NLP stack, NLG is basically that, but in reverse!

Bad:

Evvia is a Greek restaurant. Evvia is on Emerson St.
Evvia has Wi-Fi. Evvia serves wine.

Good:

Evvia, a Greek restaurant on Emerson St., has Wi-Fi and serves wine.

Bad:

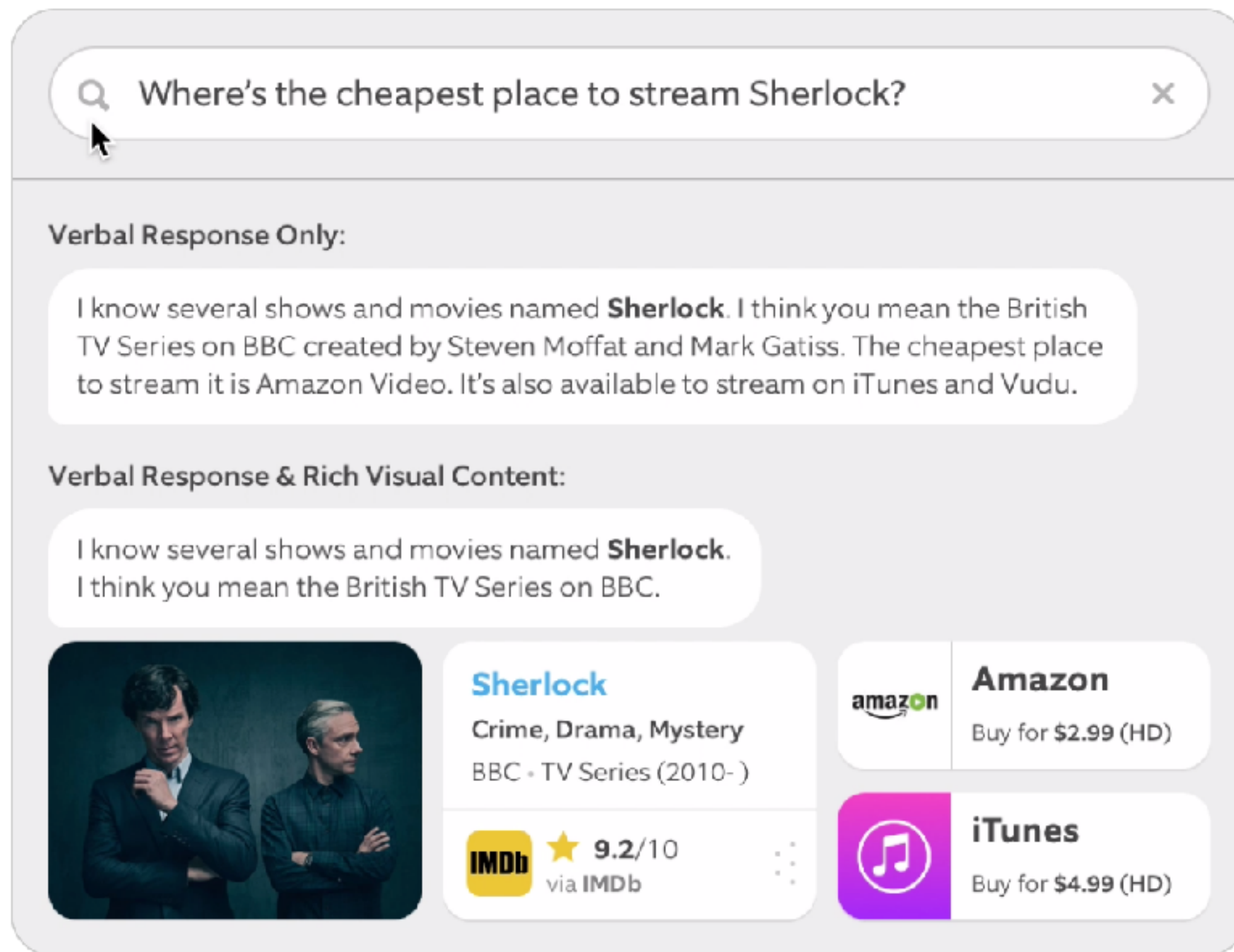
Here are a few TV shows starring Giancarlo Esposito:

Good:

I know lots of TV shows similar to Breaking Bad featuring Giancarlo Esposito, the actor that played Gus Fring:

9

Think Multi-Modal



Where's the cheapest place to stream Sherlock?

Verbal Response Only:

I know several shows and movies named **Sherlock**. I think you mean the British TV Series on BBC created by Steven Moffat and Mark Gatiss. The cheapest place to stream it is Amazon Video. It's also available to stream on iTunes and Vudu.

Verbal Response & Rich Visual Content:

I know several shows and movies named **Sherlock**. I think you mean the British TV Series on BBC.

Sherlock
Crime, Drama, Mystery
BBC • TV Series (2010-)
IMDb ★ 9.2/10 via IMDb

Amazon
Buy for \$2.99 (HD)

iTunes
Buy for \$4.99 (HD)

Users are likely to engage in a variety of modalities:

- Typing at a computer
- Speaking to a smart home hub
- Texting via a messaging app

Don't tie your responses to any one modality, build for the flexibility of rendering into a variety of formats.

And remember: speaking out a response is very different from showing the same text on screen!

Then What?

10 Measure, Improve, Measure Again

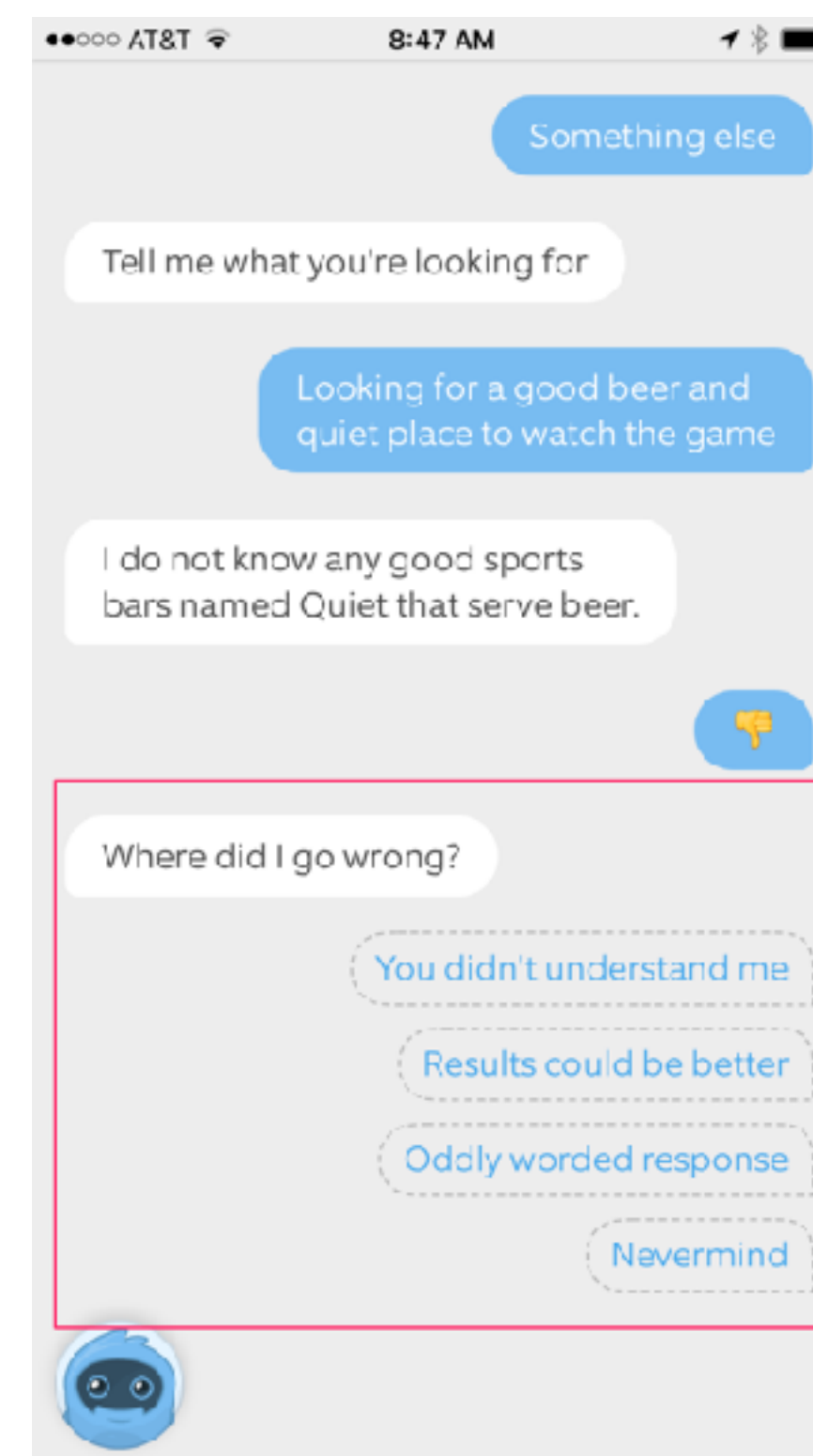
Tempting to reuse timeless search quality metrics: but do they really apply in a conversational context?

Bottom Line: How did the conversation end? Was the user pleased?

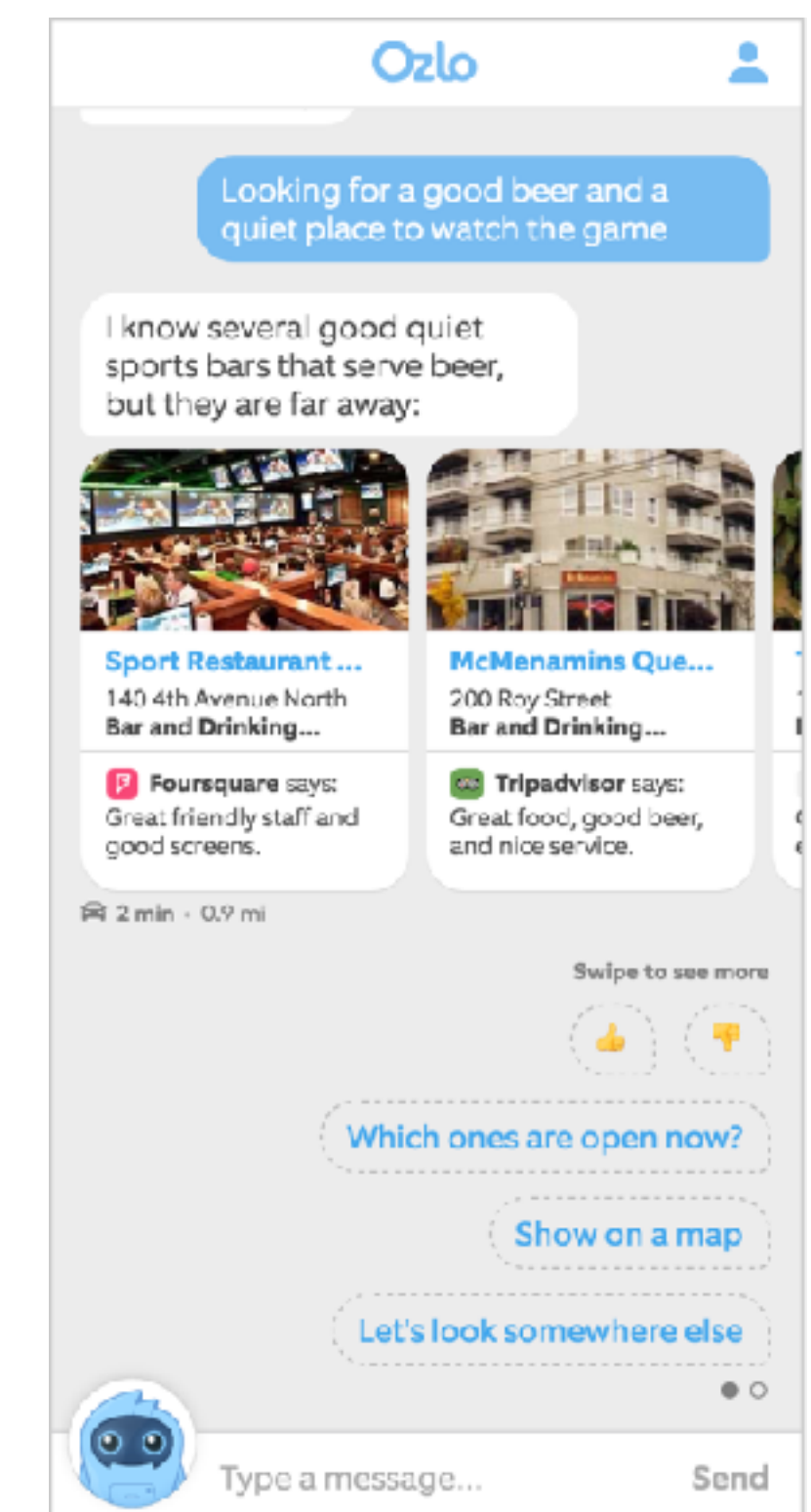
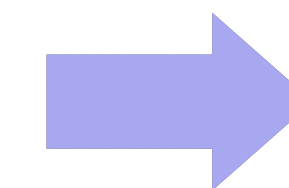
A negative conversation is exhibited by:

- Explicit negative feedback
- Highly latent
- Not well understood
- No follow on

Once you measure, you can turn this:



Into



Building an Intelligent Assistant

Know

- Model the world
- Build the right kind of data tools
- Use a hybrid text+graph search engine

Understand

- Build a real NLP stack
- Integrate “common sense” knowledge
- Separate intent from the answer

Answer

- Embrace Ambiguity
- Generate natural language responses
- Think multi-modal

Measure, Improve, Measure Again