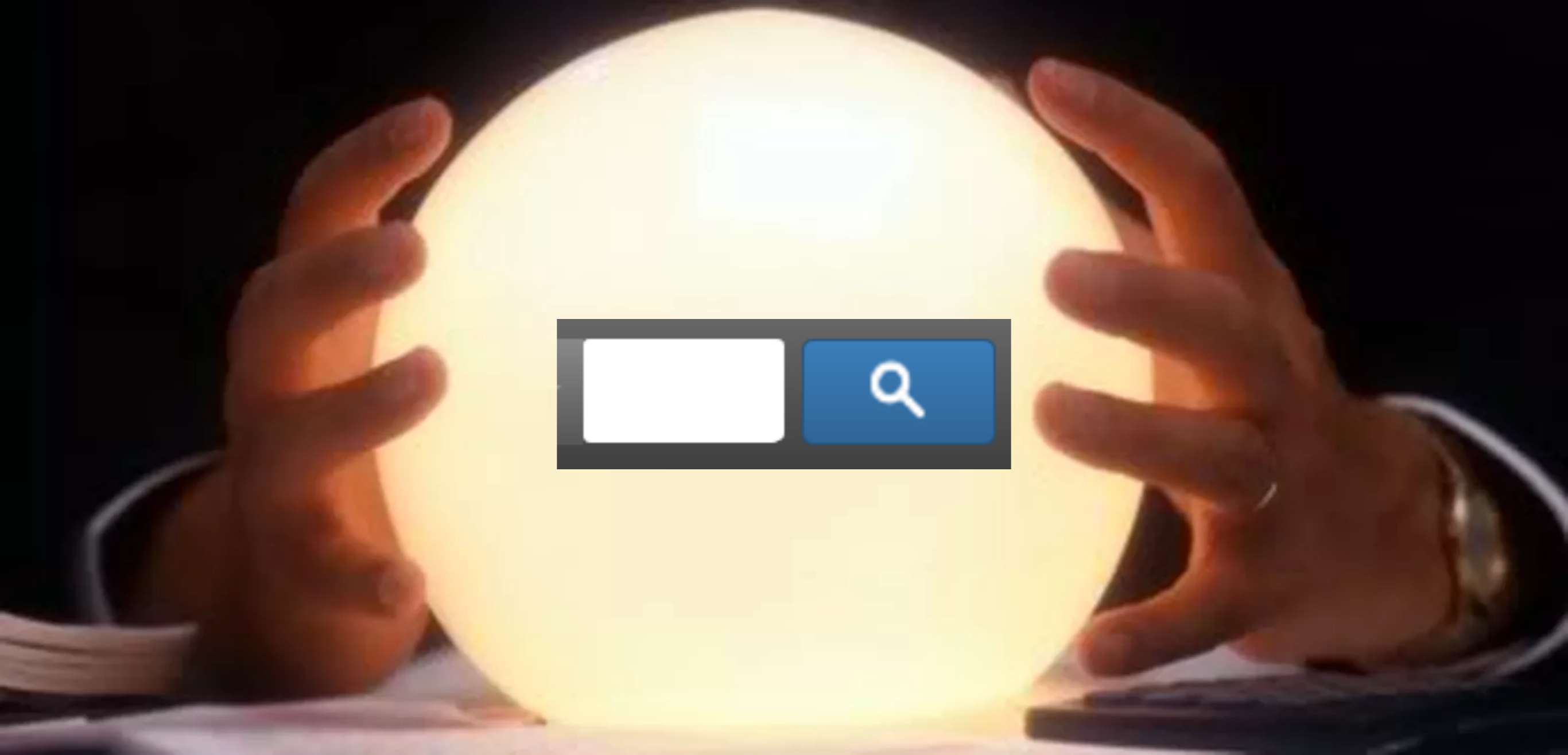


# Query Understanding: A Manifesto

**Daniel Tunkelang**

[queryunderstanding.com](http://queryunderstanding.com)



# Overview

- What is query understanding?
- Query performance prediction.
- Query rewriting.
- Query suggestions.
- Search is a conversation.


**tl;dr: Query understanding is about focusing less on the results and more on the query.**



Query understanding means  
figuring out what the searcher wants.



What query understanding is not:  
scoring and ranking of results.

The image shows four hands holding up white rectangular signs against a background of a stylized, colorful map of the world. The signs display numerical values in a large, bold, black serif font. From left to right, the values are 8.7, 9.1, 11.3, and 14.2. The hands are of various skin tones, and the background is a vibrant, abstract map with shades of blue, green, and purple.

**8.7**

**9.1**

**11.3**

**14.2**



Query understanding means not relying on ranking to filter out irrelevant results.

The screenshot shows the LinkedIn search interface. At the top, the LinkedIn logo and 'PREMIUM' are visible. The search bar contains 'evan pinterest' with a search icon to the right. Below the search bar, navigation tabs for 'Home', 'Profile', 'My Network', 'Jobs', and 'Interests' are present. The search results section is titled '2 results for evan pinterest'. A red-bordered box highlights a message: 'Some search results have been filtered to improve relevance. Show all results'. Below this, two search results are shown. The first result is for 'Evan Sharp', a 2nd-degree connection, Founder at Pinterest, located in the San Francisco Bay Area, with 21 shared connections. The second result is for 'Evan Jones', also a 2nd-degree connection, SRE Manager at Pinterest, located in the San Francisco Bay Area, with 5 shared connections. On the left side, there are filters for 'All', 'People', 'Jobs', 'Companies', 'Groups', 'Universities', 'Posts', and 'Inbox'. Below these, a 'Relationship' section is visible with checkboxes for 'All', '1st Connections', and '2nd Connections', where 'All' is selected.

**Search**

[Advanced >](#)

**2 results for evan pinterest**

Some search results have been filtered to improve relevance.  
[Show all results](#)

**All**

- People
- Jobs
- Companies
- Groups
- Universities
- Posts
- Inbox

**Relationship**

- All
- 1st Connections
- 2nd Connections

**Evan Sharp** **2nd**  
Founder at **Pinterest**  
San Francisco Bay Area • Design  
▶ **21 shared connections** • [Similar](#)

Current: Founder at **Pinterest**  
Past: Product Designer at Facebook  
Education: The University of Chicago  
Columbia University - Graduate School of Architecture, Planning a...

**Evan Jones** **2nd**  
SRE Manager at **Pinterest**  
San Francisco Bay Area • Internet  
▶ **5 shared connections** • [Similar](#)

Current: SRE Manager at **Pinterest**

Challenge: accurately identify searcher's intent while minimizing searcher's effort.



# Old-school search: searcher enters query, search engine interprets query.



kittens



All

Images

Videos

Shopping

News

More ▾

Search tools

About 48,000,000 results (0.56 seconds)

Images for kittens

Report images



More images for kittens

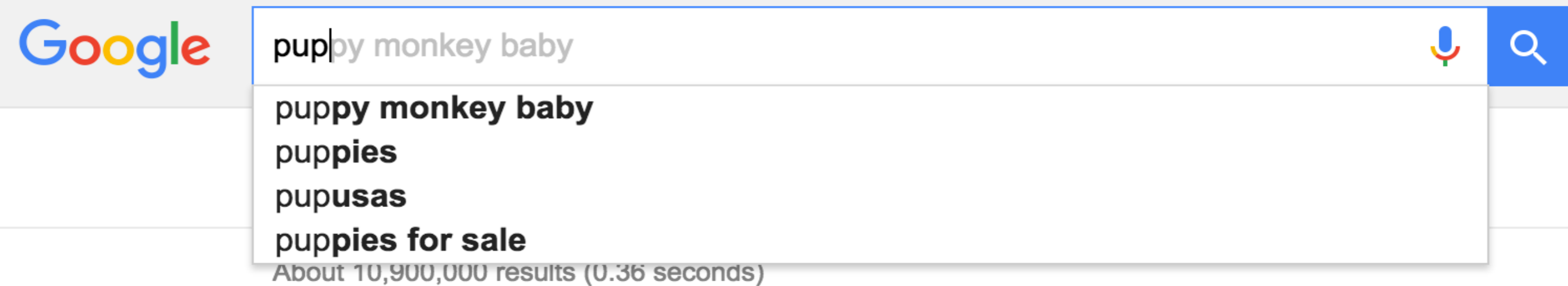
[Kitten - Wikipedia, the free encyclopedia](https://en.wikipedia.org/wiki/Kitten)

<https://en.wikipedia.org/wiki/Kitten> ▾ Wikipedia ▾

A **kitten** is a juvenile cat. After being born, **kittens** are totally dependent on their mother for survival and they do not normally open their eyes until after seven to ...



Modern search:  
searcher starts to type, search engine suggests queries.



Google puppy monkey baby

- puppy monkey baby
- puppies
- pupusas
- puppies for sale

About 10,900,000 results (0.36 seconds)

Images for puppy monkey baby

Report images



More images for puppy monkey baby

Mtn Dew Kickstart: Puppymonkeybaby | Super Bowl Spot - YouTube

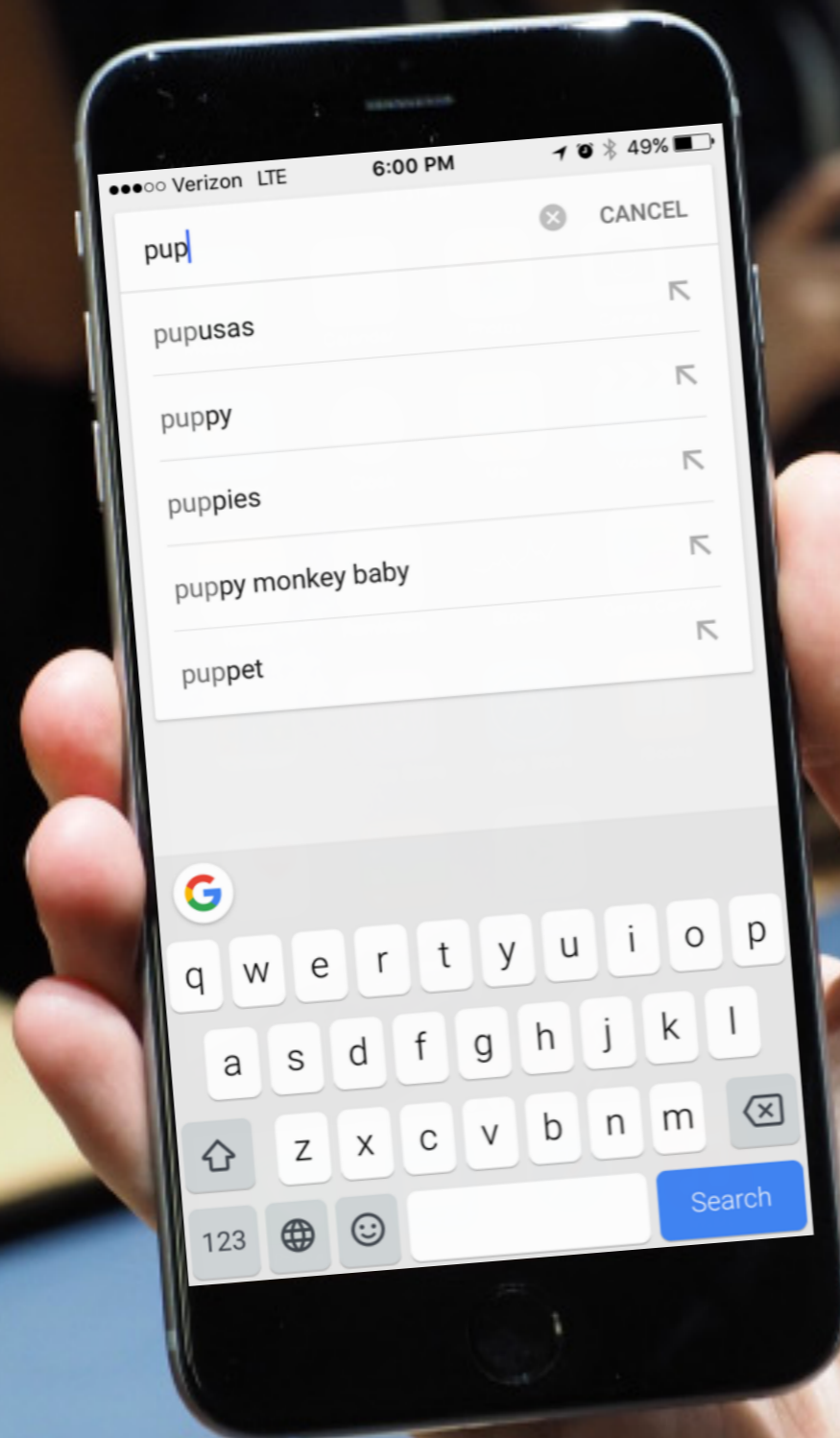


<https://www.youtube.com/watch?v=ql7uY36-LwA>

Feb 3, 2016 - Uploaded by Mountain Dew

When three friends decide to stay home for the night, a surprise visitor arrives with MTN DEW Kickstart on a ...

And in 2016,  
modern search means mobile search.





March Hare: "You should say what you mean."  
Alice: "I do. At least I mean what I say."





# But what if the search engine doesn't know what the searcher means?



**WIKIPEDIA**  
The Free Encyclopedia

Not logged in [Talk](#) [Contributions](#) [Create account](#) [Log in](#)

Article [Talk](#)

Read

[Edit](#)

[View history](#)

Search



## Springfield

From Wikipedia, the free encyclopedia

**Springfield** may refer to:

- [Springfield \(toponym\)](#), the place name in general

### Places and locations [[edit](#)]

#### **Australia** [[edit](#)]

- [Springfield, New South Wales \(Central Coast\)](#)
- [Springfield, New South Wales \(Snowy Monaro Regional Council\)](#)
- [Springfield, Queensland](#)
- [Springfield, South Australia](#)
- [Springfield, Victoria](#) in Buloke Shire, north-western Victoria
- [Springfield, Victoria \(Macedon Ranges\)](#) in central Victoria

#### **Belize** [[edit](#)]

- [Springfield, Belize](#)

#### **Canada** [[edit](#)]

- [Rural Municipality of Springfield](#) in Manitoba
  - [Springfield \(provincial electoral district\)](#), an electoral division in Manitoba

a multilingual tree encyclopedia  
**Wiktionary**  
['wɪkʃənəri] *n.*,  
a wiki-based Open  
Content dictionary  
Wikipedia [wɪˈpi:di.ə]

Look up ***Springfield*** in Wiktionary, the free dictionary.

#### **Contents** [[hide](#)]

- 1 Places and locations
  - 1.1 Australia
  - 1.2 Belize
  - 1.3 Canada
  - 1.4 Ireland
  - 1.5 New Zealand
  - 1.6 South Africa
  - 1.7 United Kingdom
  - 1.8 United States
  - 1.9 Amtrak stations
  - 1.10 Fictional places
- 2 People
  - 2.1 Fictional characters
- 3 Other
- 4 See also

[Main page](#)

[Contents](#)

[Featured content](#)

[Current events](#)

[Random article](#)

[Donate to Wikipedia](#)

[Wikipedia store](#)

Interaction

[Help](#)

[About Wikipedia](#)

[Community portal](#)

[Recent changes](#)

[Contact page](#)

Tools

[What links here](#)

[Related changes](#)

[Upload file](#)

[Special pages](#)

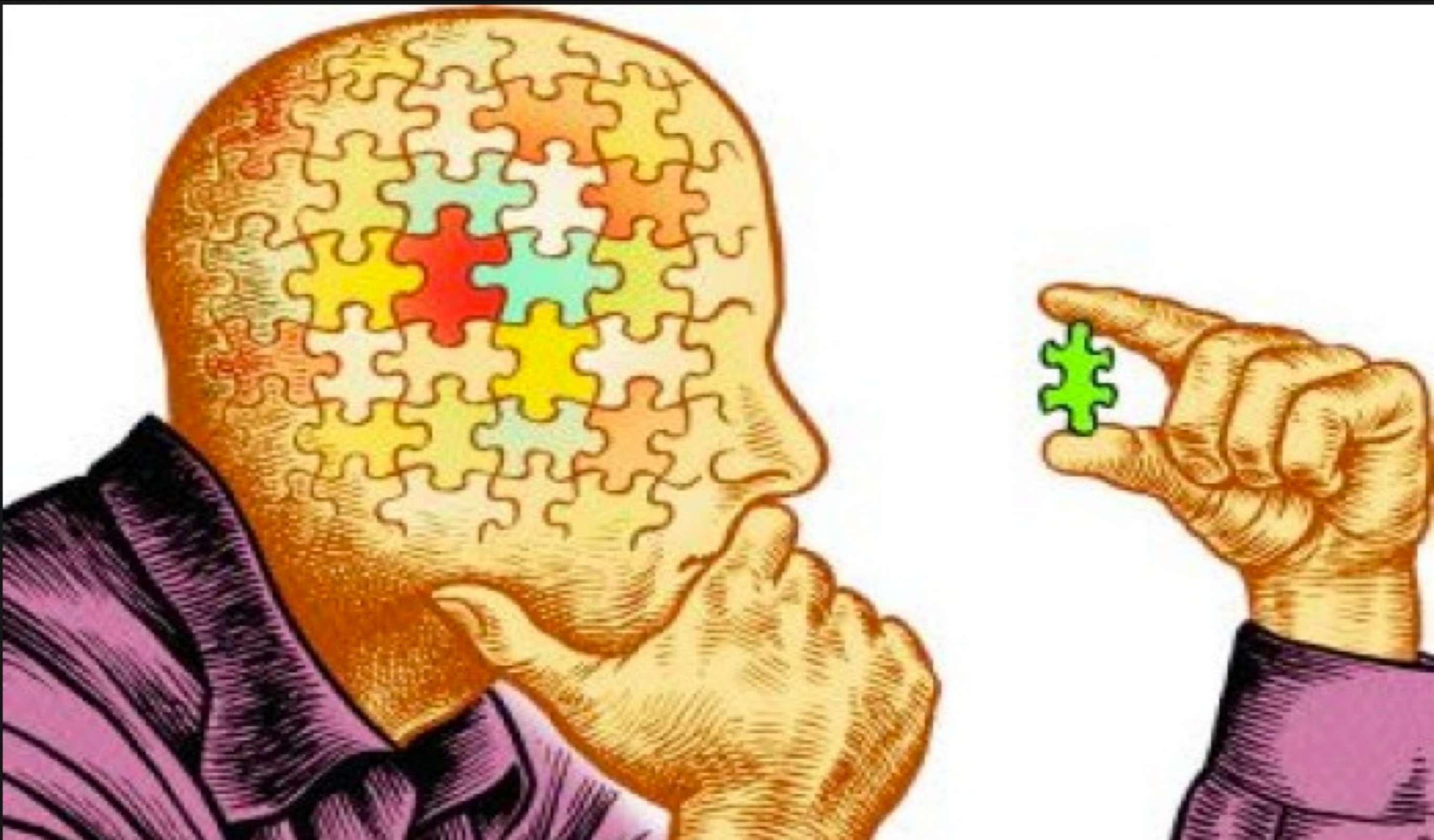
[Permanent link](#)

[Page information](#)

[Wikidata item](#)

[Cite this page](#)

And how does a search engine achieve the self-awareness to know what it doesn't know?





Query understanding is about measuring and optimizing for query performance.





History is the best teacher.  
For common queries, learn from historical performance.





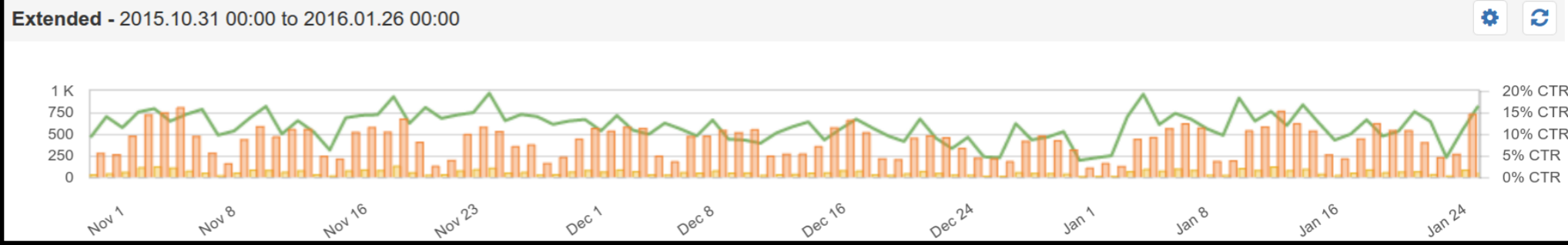
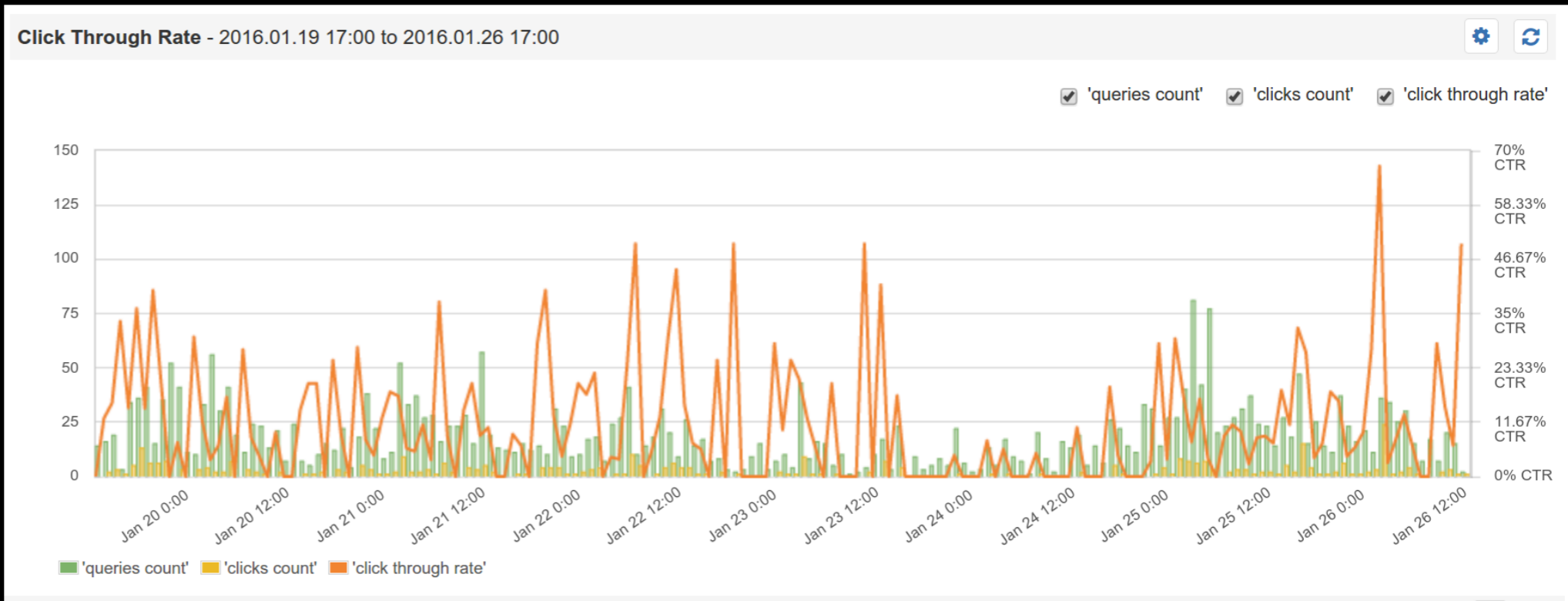
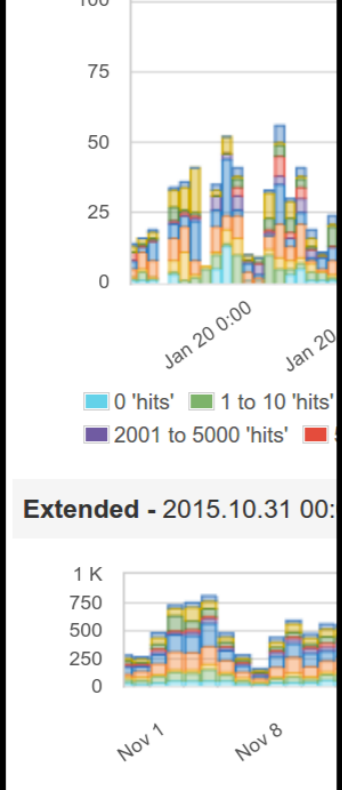
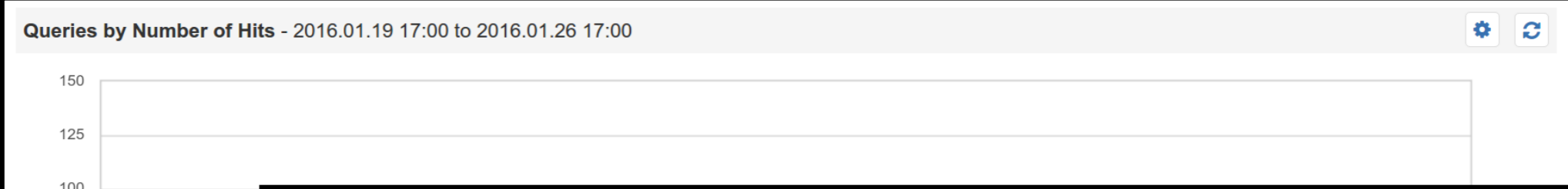
# How you define performance depends on why searchers use your site. But define it and measure it!

By 'Queries Count' - 01/19/2016 16:00 - 01/26/2016 16:00

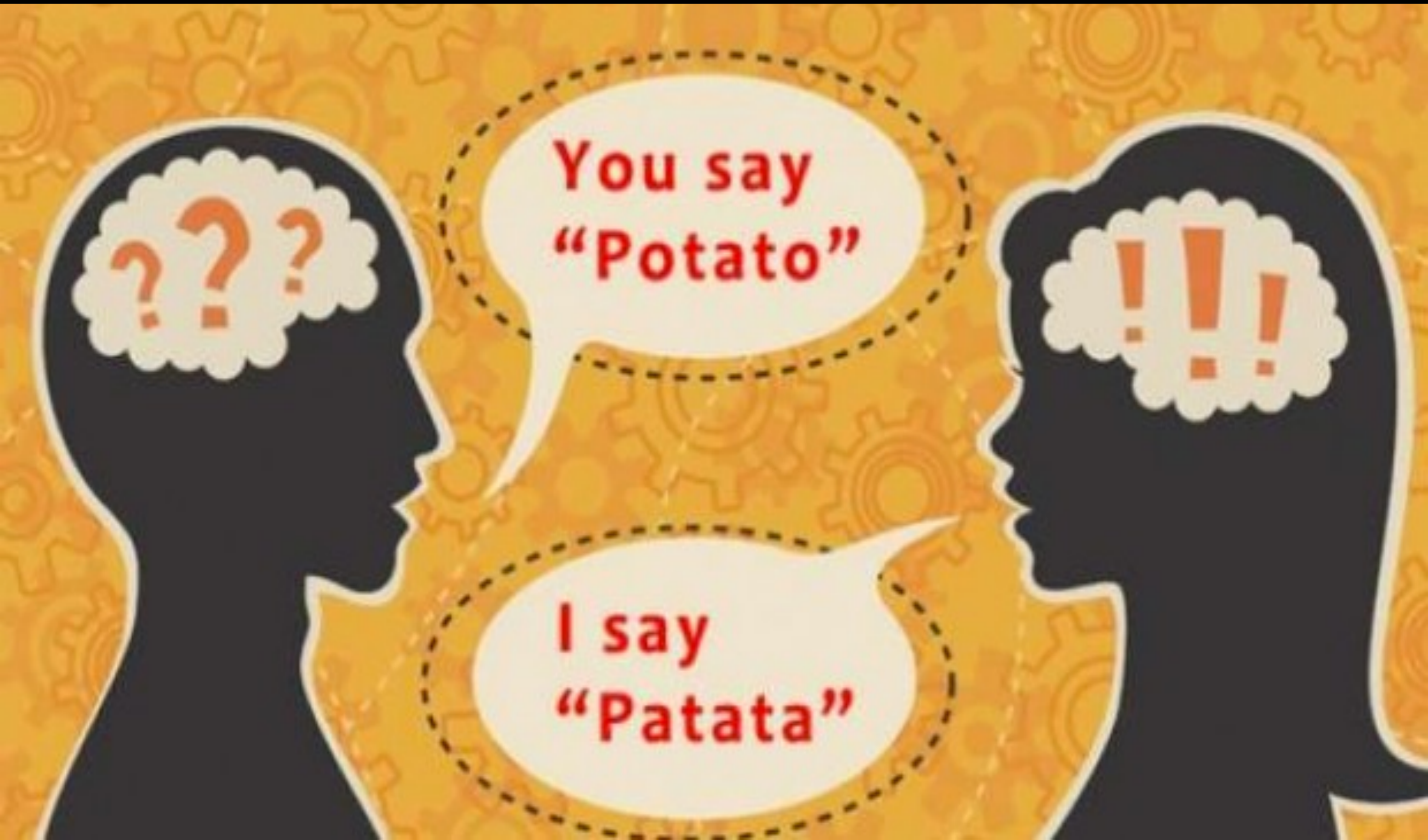
Totals: **Queries**1040 **Clicks**121 **Queries without clicks**919 **CTR**11.63 **Zero hits**11 **Zero hits % of total**1.06 **MRR**0.45

Show/hide: Select All Counts, Hits, Clicks Latencies Zero Hits Search: Show 50

- Query String
- hdfs
- hbase
- mapreduce
- license
- mysql
- pig
- +%5C%22...22you
- sqoop
- Search with Apache



Can also learn from similar queries:  
re-ordered terms, subqueries, similar results, etc.



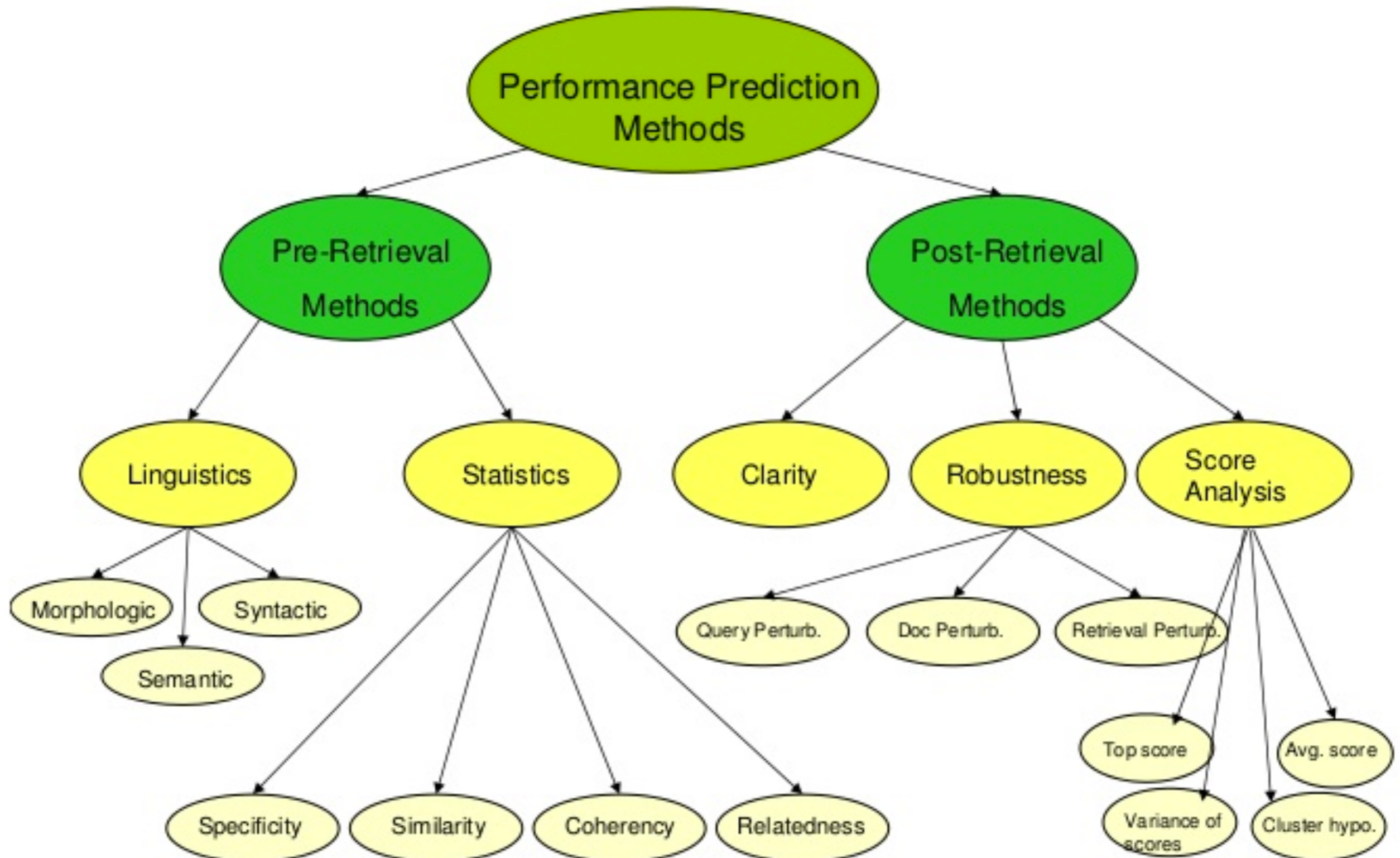


Stuck in the long tail?  
Predict performance from query and result set features.





Lots of academic work on query performance prediction.









# Rewriting to improve precision: query segmentation, field restriction, query classification.

**verticals:**  
people, jobs

**intent:** exploratory



**tag:** skill OR title  
**related skills:**  
search, ranking, ...

**tag:** company  
**id:** 1337  
**industry:** internet





Rewriting to improve recall:  
query expansion, query relaxation.





When query expansion is done right, users barely notice.  
Query relaxation is riskier and should be messaged.

sf subway schedule

All News Maps Images Shopping More Search tools

About 22,100,000 results (0.45 seconds)

[Schedules & Trip Planners | SFMTA - San Francisco Municipal ...](#)  
<https://www.sfmta.com/node/65376> San Francisco Mu  
Trip planners and timetables to get you where you want to go

[Routes & Stops | SFMTA - San Francisco M](#)  
<https://www.sfmta.com/node/71396> San Francisco Mu  
KT: K-Ingleside/T-Third Street · **Schedule** · Live Data · Rece  
Receive Alerts · L-Taraval · **Schedule** · Live Data · Receive

[Muni Transit | SFMTA - San Francisco Muni](#)  
<https://www.sfmta.com/node/65377> San Francisco Mu  
LED sign inside bus shelter showing time and **schedule** infor  
riding Muni buses, our **Metro** light rail, and unforgettable ...

[Bay Area Rapid Transit | bart.gov](#)  
<https://www.bart.gov/> Bay Area Rapid Transit  
**Schedules** · Overview · Real Time Departures · Service Adv  
BART Map · Close **BART** Map. Departing from? Station.

query performance prediction for ecommerce search

All News Images Videos Shopping More Search tools

About 23,000,000 results (0.71 seconds)

[\[PDF\] Query Performance Prediction in Web Search Environments](#)  
[maroo.cs.umass.edu/getpdf.php?id=726](http://maroo.cs.umass.edu/getpdf.php?id=726)  
by Y Zhou - Cited by 149 - Related articles  
types of queries in web search environments: content-based and ... **Query performance prediction**  
has many applications in a variety of information retrieval (IR) ...

Missing: ~~ecommerce~~

[Query performance prediction in web search ... - ACM Digital Library](#)  
[dl.acm.org/citation.cfm?id=1277835](http://dl.acm.org/citation.cfm?id=1277835) Association for Computing Machinery  
by Y Zhou - 2007 - Cited by 149 - Related articles  
Jul 23, 2007 - **Query performance prediction** in web search environments .... Deolalikar , Kang Li,  
Atypical Queries in eCommerce, Proceedings of the 24th ...

Query rewriting should improve query performance.  
 If it doesn't, you're doing wrong. Measure!

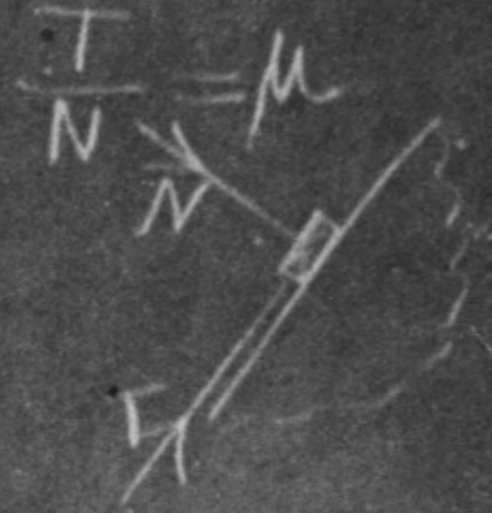
1095	372.8	410.6
955	281.0	270.0
8599	281.0	270.0
758	270.0	
609		
401		
228		
614	244.4	

Velocity of elasticity waves in kilometers per second

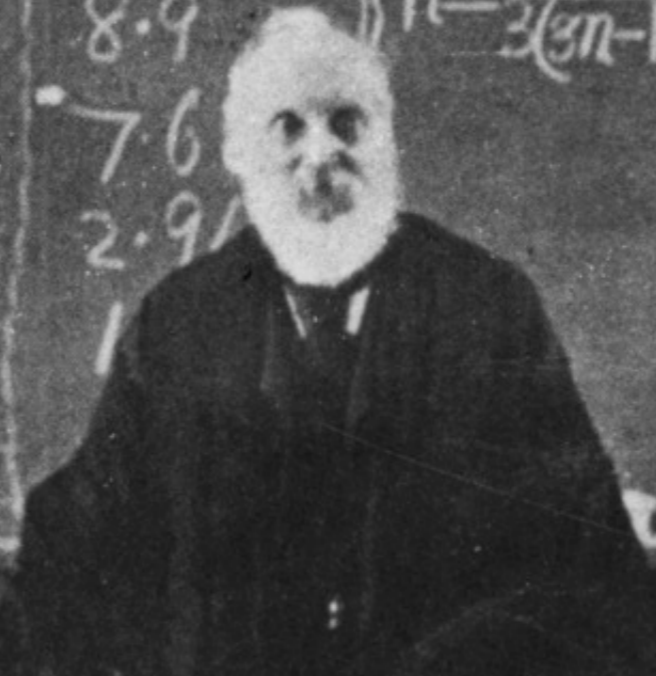
Longitudinal in fluid $v = \sqrt{\frac{K}{\rho}}$	Distortional $v = \sqrt{\frac{B}{\rho}}$	Condensation $v = \sqrt{\frac{K + \frac{4}{3}n}{\rho}}$	Transversional $v = \sqrt{\frac{n}{\rho}}$
Copper	3.72	2.24	5.01
Iron	5.11	3.17	5.92
Glass	4.53	2.85	5.02
Water			1.49
Air			1

$$M = \frac{9nR}{3R + n}$$

$$R = \frac{Mn}{3(3n - M)}$$



$$\frac{T}{N} < \mu$$





Rewriting is great when searchers type whole queries.  
But what about autocomplete and search suggestions?

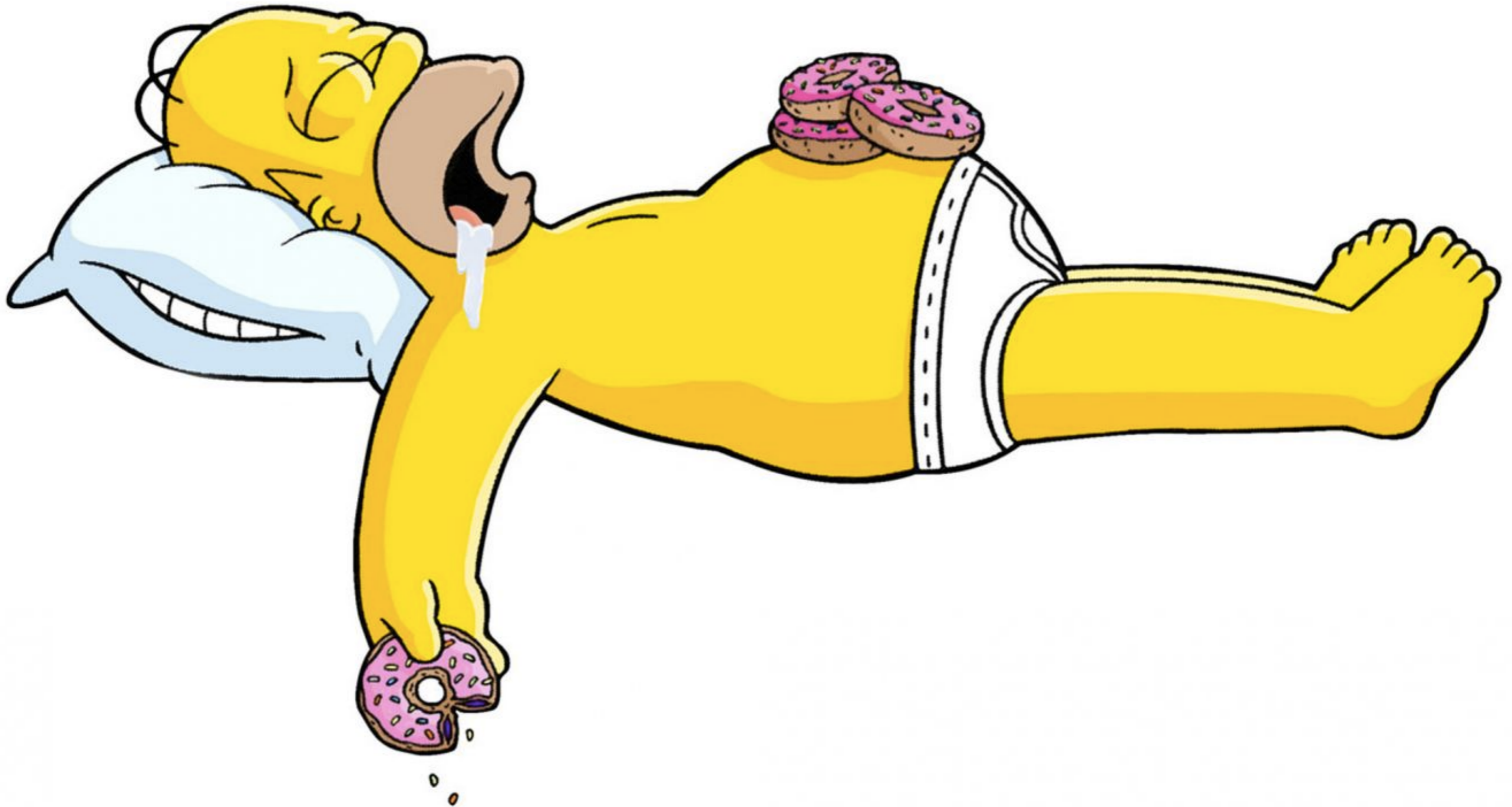




Query suggestions have two jobs:  
reduce searcher effort and improve query performance.

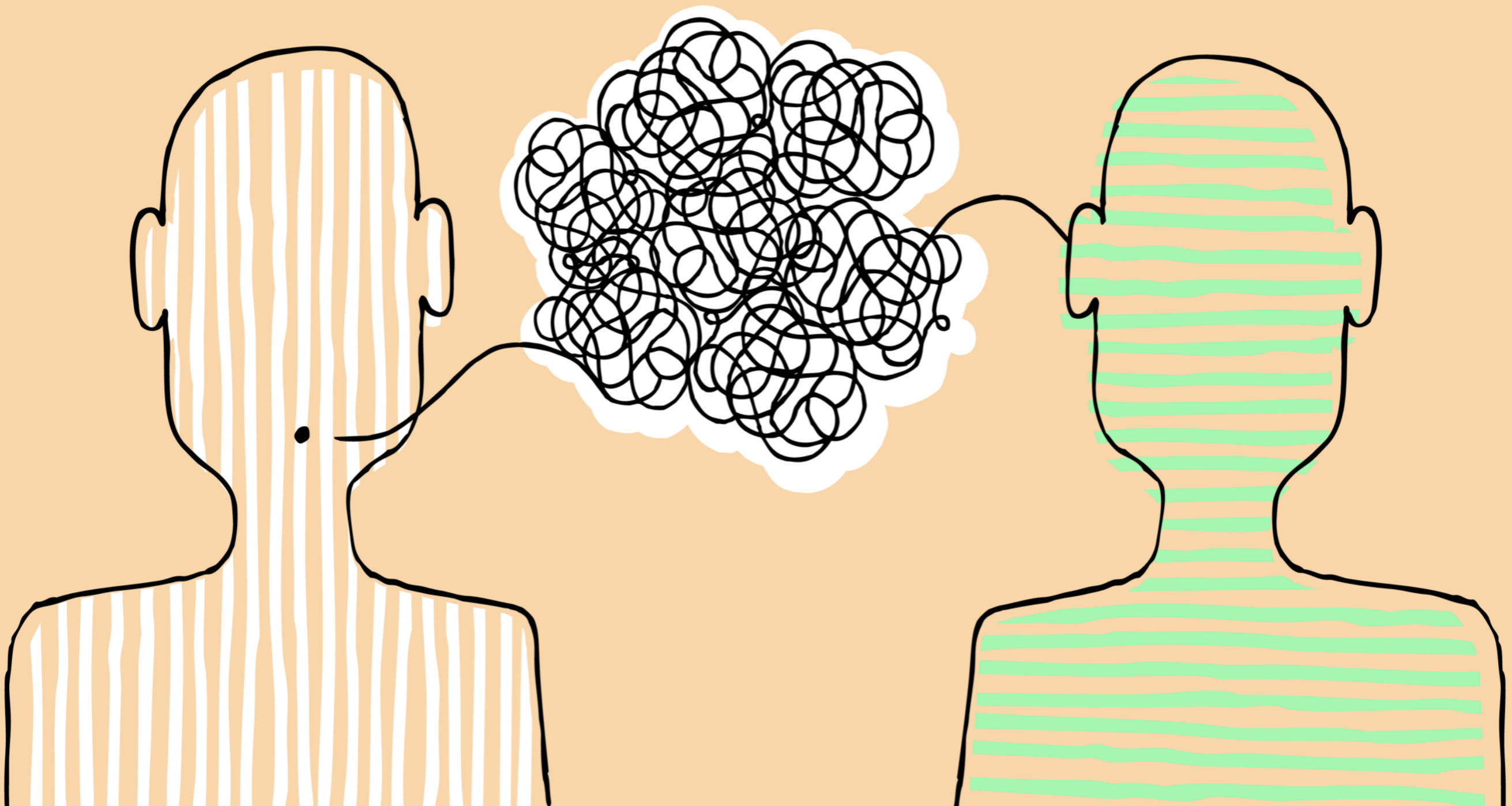


Searchers are people. People are lazy.  
All else equal, help them type as little as possible.





But don't let searchers be too lazy!  
It's easy to create miscommunication.



# Search suggestions offer strong information scent. But searcher may not recognize query ambiguity.

Knots

Gagged

Clip

Dress

Bar

Bow

Men's

To Bed

Blanket



**\$95.00 - \$100.00**

Twist & Tie your swimsuit any way you want it! Modest one piece swimsuit in bright hawaiian print! Bev Swim

from [beverlyswimwear.com](#)  
2.3k

Nikki Brees  
CLOTHING



**\$19.00**

from Pursuit

**Grograin Solid Slim Green Teal Tie**

Grograin Solid Slim Green Teal Tie  
2

Pinned by  
Pursuit



**\$25.00**

from Play\*Patch

**Dr Who Tie**

Dr Who Tie  
247

Play\*Patch  
Ties



**\$54.95**

from BuffaloJackson.com

**Bow Tie - Trout - Rainbow Trout Pink**

Trout Bow Tie (Rainbow Trout Pink)  
169

Well Prepped  
Mens Prep Essentials



**\$35.99**

from Moooh!!

**Inspired Venice Tie**

Inspired Venice Tie  
4

Pinned by  
Moooh!!



Reduce the likelihood of a bait-and-switch:  
only offer high-performance search suggestions.



Should you score search suggestions based on total engagement or query performance?





Total engagement is the value delivered by search.  
But low query performance means unhappy users.



**MONEY**

**V.S.**



**HAPPINESS**

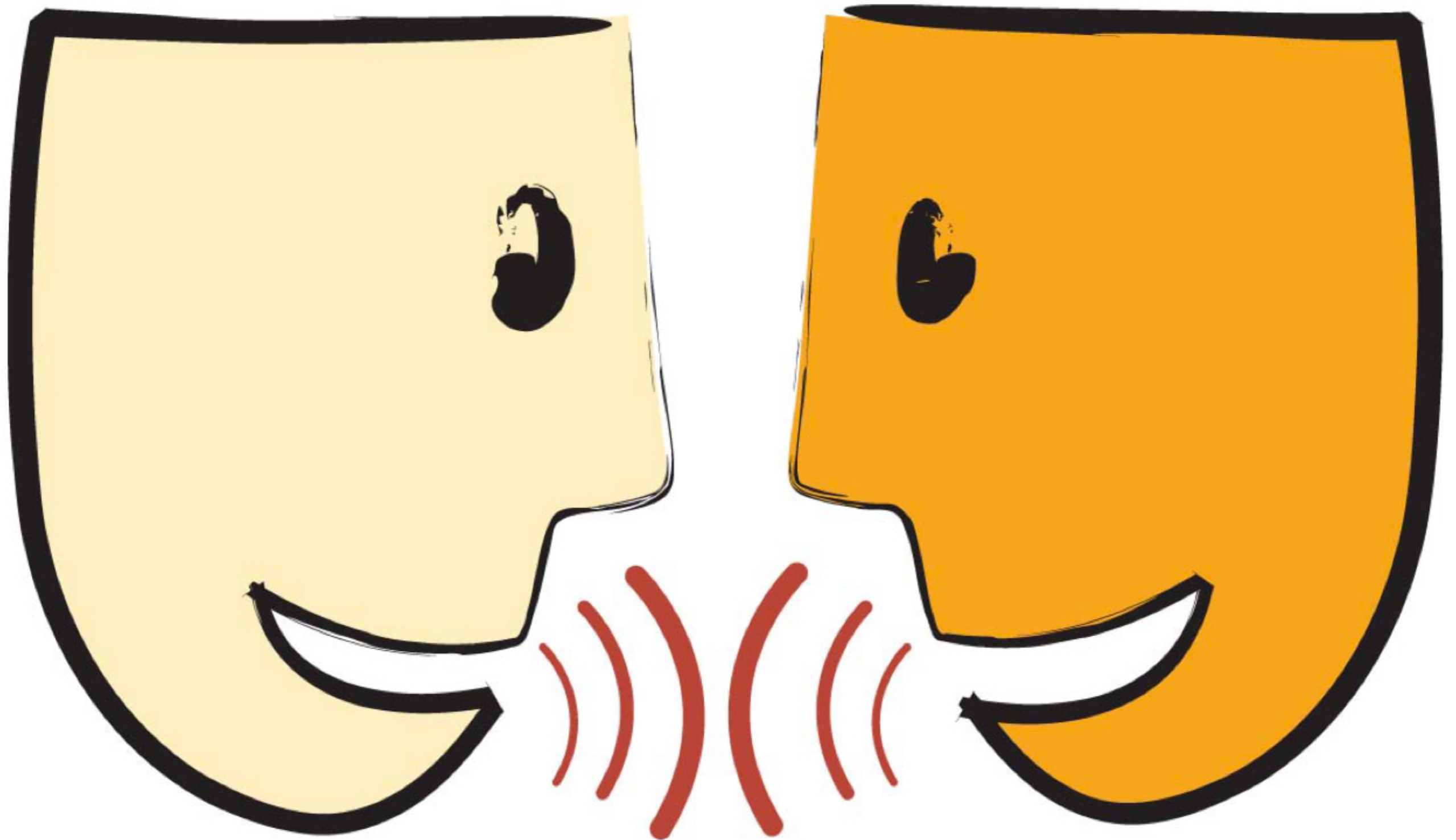


Can combine as single utility function. Or score using total engagement but require minimum query performance.



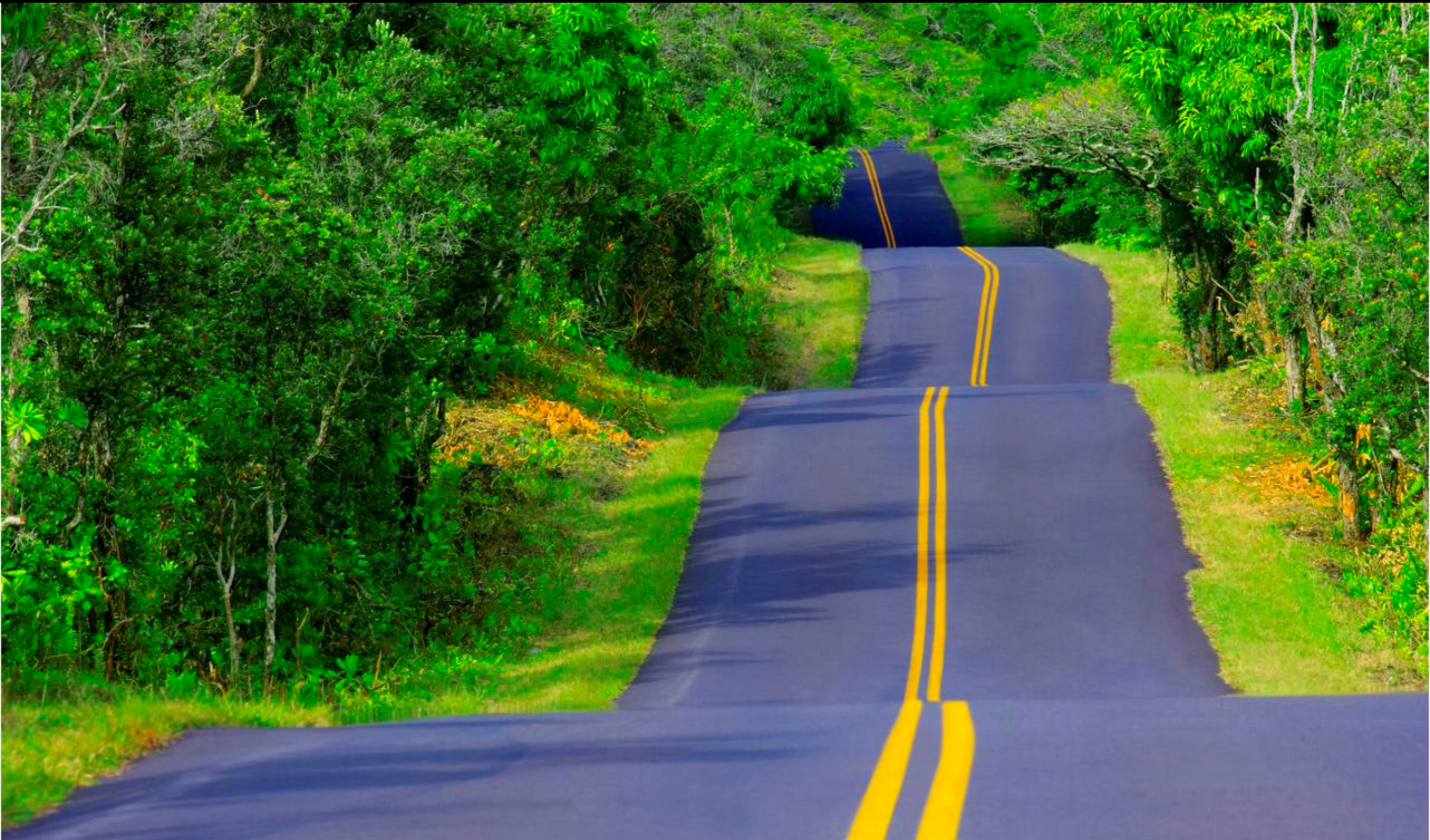


Finally, let's remember that search is more than a single query and response. It's a conversation.





Don't focus so much on single-query performance that you forget to optimize for the overall journey.





The best journey is a sequence of small successes.  
Not a sequence of frustrations followed by success.



# Summary

- Query understanding means figuring out what the searcher wants.
- Measure and optimize for query performance.
- Use query rewriting to improve precision and recall, in order to improve query performance.
- Query suggestions have two jobs: reduce searcher effort and improve query performance.
- Search is a conversation.

**tl;dr: Query understanding is about focusing less on the results and more on the query.**



Thank you! Learn more at [queryunderstanding.com](http://queryunderstanding.com)!

