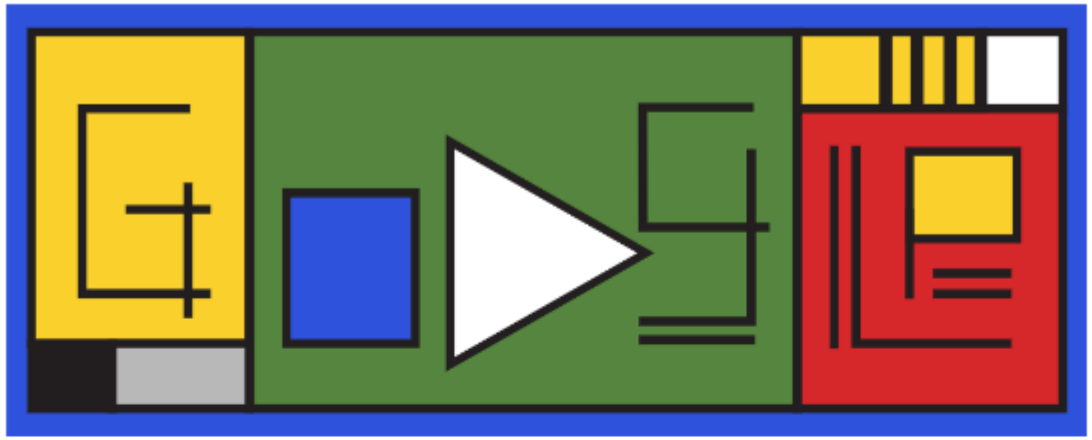


Google

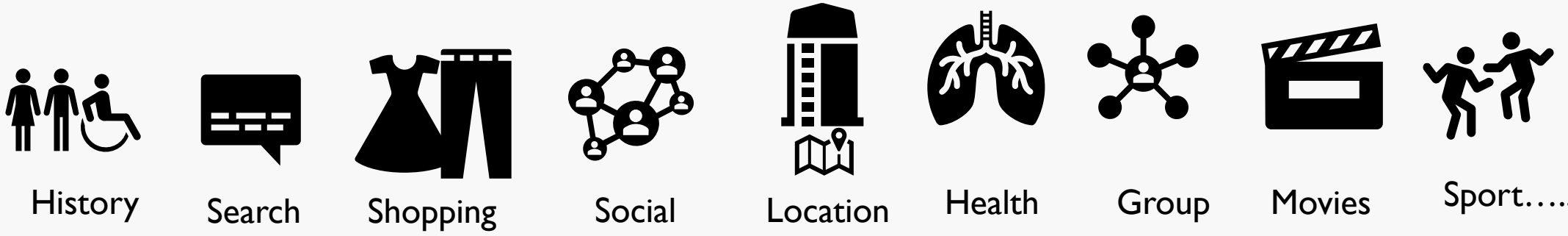
google **search recommendation algorithm** Remove
google **search recommendation** Remove
google
google **translate**
google **drive**
google **maps**
google **docs**
google **classroom**
google **flights**
google **calendar**

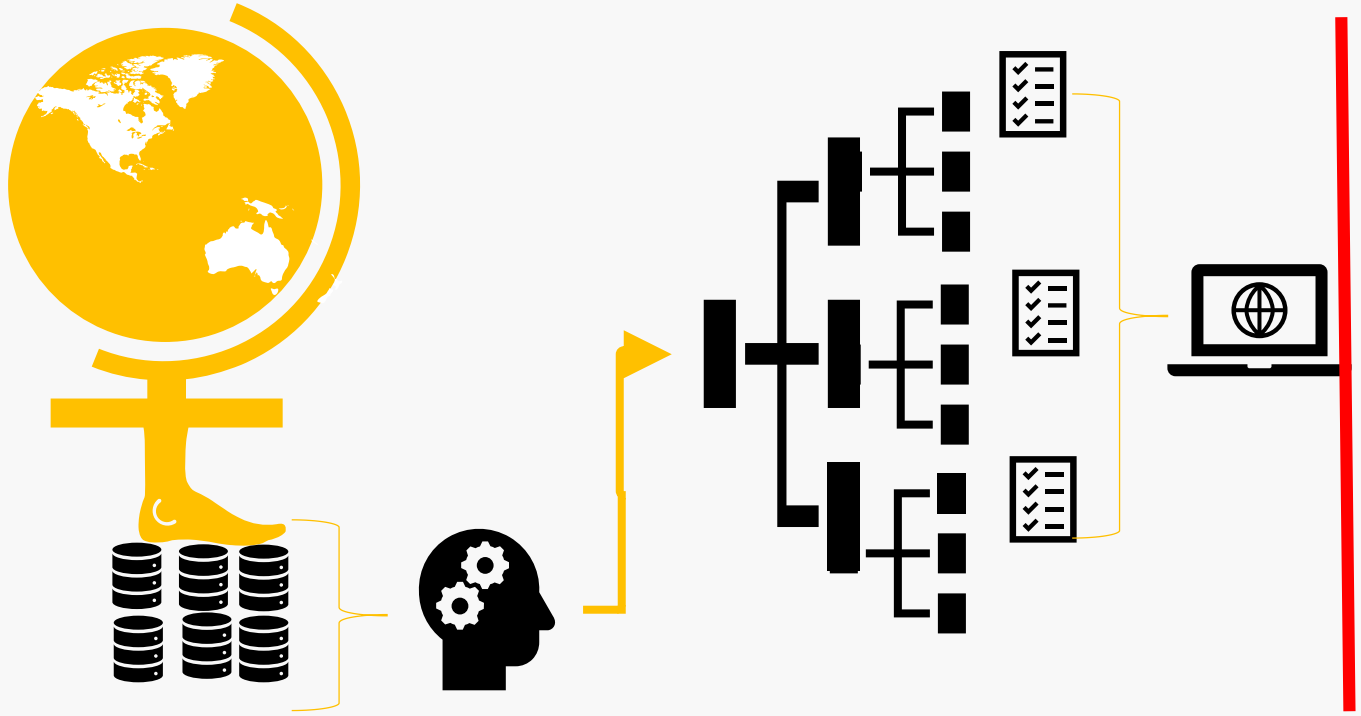
Google Search I'm Feeling Lucky

Report inappropriate predictions

GOOGLE SEARCH RECOMMENDATION

GOOGLE SEARCH RECOMMENDATION





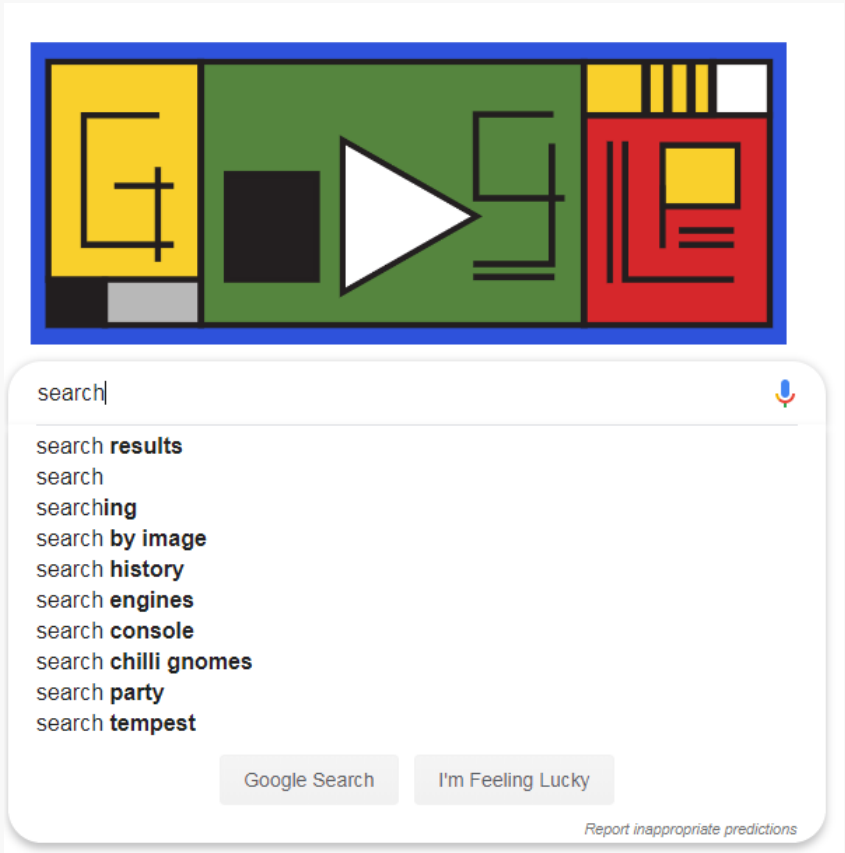
We give lots of info to google while we use any of the services offered by Google.

Google is learning the info and tailoring the info as per best of our interest or use.

There is hybrid approach behind this where google is able to respond to our query.

SEARCH PREDICATION BY LANGUAGE

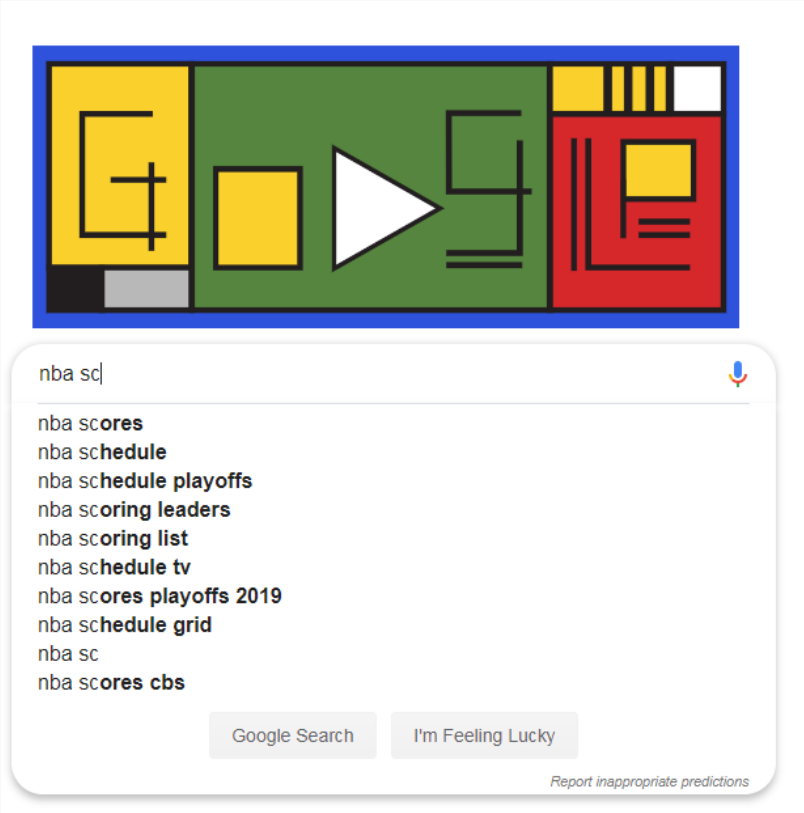
English



Hindi

Based on keyed text's language Google is able to decide on the search predication language.

GIVING YOU FRESHER, MORE RECENT SEARCH RESULTS



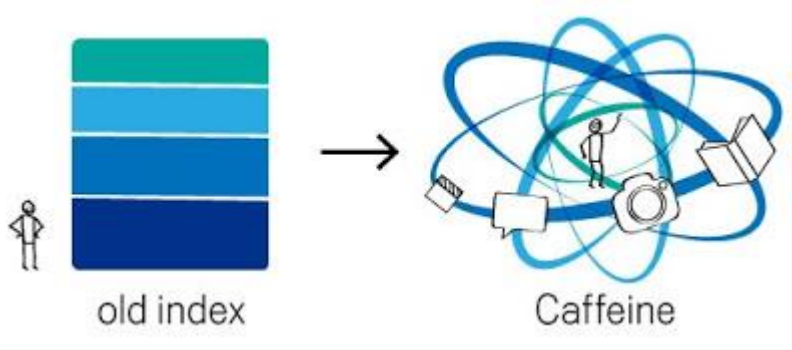
Search results, like warm cookies right out of the oven or cool refreshing fruit on a hot summer's day, are best when they're fresh. Even if you don't specify it in your search, you probably want search results that are relevant and recent.

Based on keyed text's language Google is able to decide on the search predication language.

GIVING YOU FRESHER, MORE RECENT SEARCH RESULTS

Old index < 2010

- Several layers
- Very few were refreshed at a faster rate than others
- The main layer would update every couple of weeks.
- To refresh a layer of the old index, Google would analyze the entire web, which meant there was a significant delay between when Content was created vs when google found it.



With Caffeine and Panda After 2010

- Google has categorized data for search result :
  - Regularly recurring events:
  - Frequent updates:
  - Recent events or hot topics:

Caffeine lets Google index web pages on an enormous scale. In fact, every second Caffeine processes hundreds of thousands of pages in parallel. If this were a pile of paper it would grow three miles taller every second

SCENARIO DESIGN

- Who are your target customers?
  - Any user who is trying to find some info from web.
  - This include big segment of user base, organization , education , Government and many more.

- What are their key goals?
  - User's searching on Google should get most relevant information.
  - It should be most accurate, as per the trend.
  - User should be able to find most relevant info without scrolling too much.
    That's why we see lots competition for ranking on 1st page of google . for
  - Able to propose /predict possible search term in case of ambiguity in first search
- How can Google help their users accomplish those goals?
  - Google is using traditional / Machine Learning both content based and Collaborative to predicts and recommend search terms

SCENARIO DESIGN : REVERSE ENGINEERING

- With the help of Machine Learning and help of PANDA rollout of google search engine google tries to read some key information from the website as part data collection .

Google Panda targets websites with the following problems:

•Duplicate content
•Low quality and thin pages that provide no real value to users
•Websites with a bad user experience
•External links to bad websites
•Ad placement
•Word Count & On Page Factors
•Advertising & Affiliates

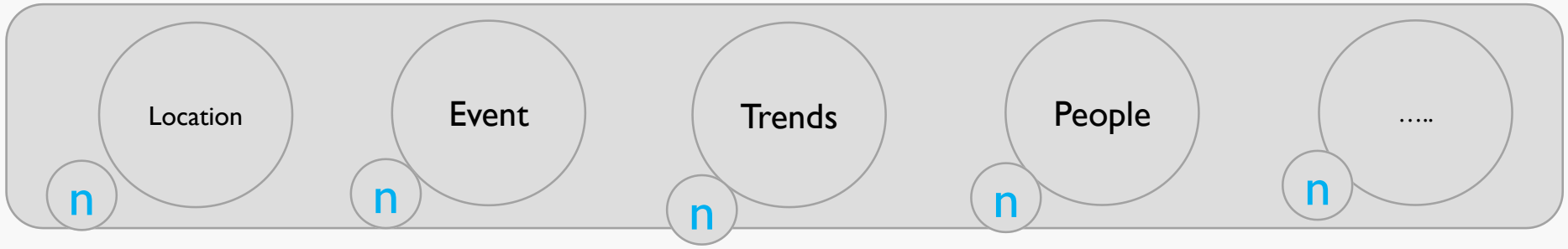
•backlinks profile
•How with other site
•API calls
•Load time
•Sites with too many advertisements
•Content from sites
•Poor grammar and spelling
•Too many broken links
•Comments

http://www.bayardo.org/ps/vldb2009.pdf

SCENARIO DESIGN : SEARCH TERM

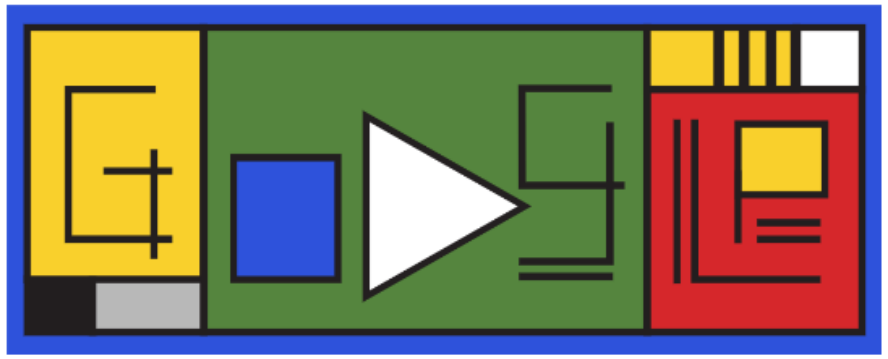
Query <Wait >

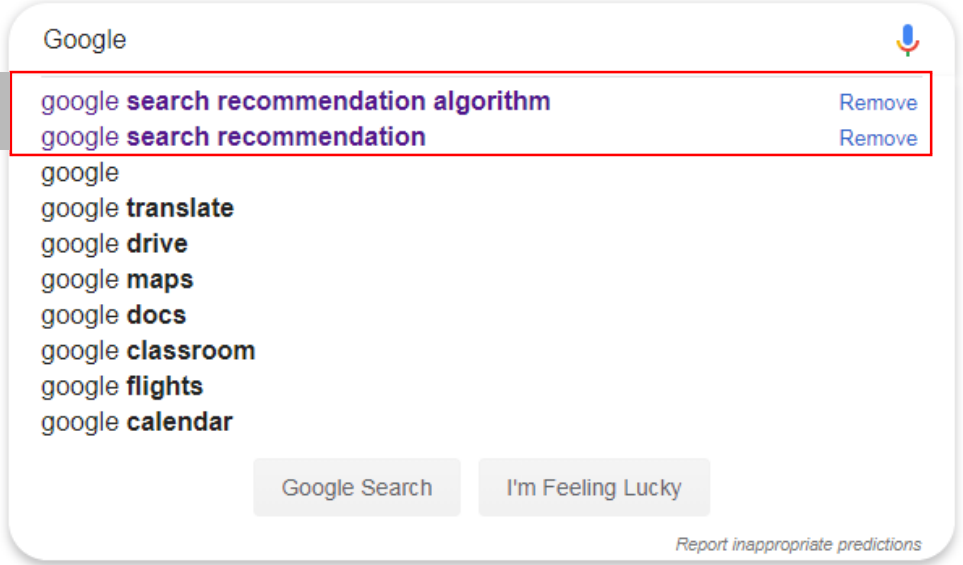
Specific queries. E.g., "Does Netscape support the JDK 1.1 code-signing API?"
Broad-topic queries. E.g., "Find information about the Java programming language."
Similar-page queries. E.g., "Find pages 'similar' to java.sun.com."



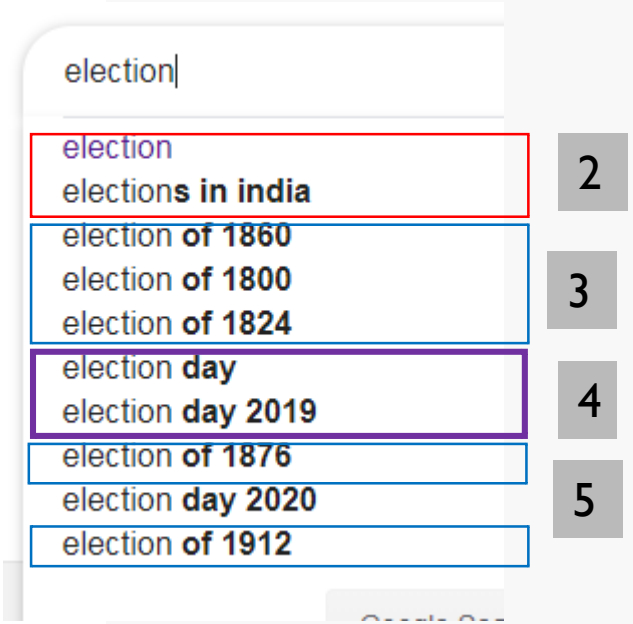
-Getting related terms from the cluster of data, based on users local Search history and Search based on Geolocation, Social Trends, current events, and many more.
-Giving some score to each terms based on the freshness of the data and then listing them users sections as user types .
- Google Search predication also resorve some section for some part of data. (next page)

SEARCH TERMS CATEGORIZATION





1



2 3 4 5

1 Data from searched History.

2 Data from Trending Events globally

5 Data from future events.

3 Data from Historical Event.

4 Data from Possible current years events

USING TERMS

Danny Sullivan ✓
@dannysullivan

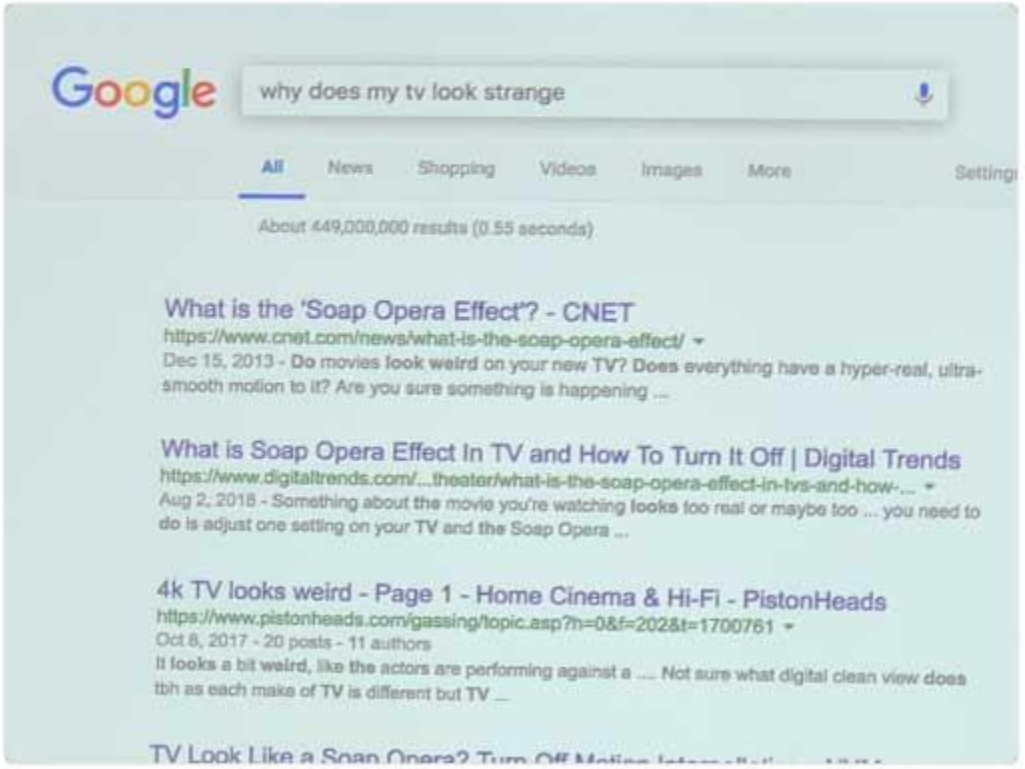
Following

This is a look back at a big change in search but which continues to be important: understanding synonyms. How people search is often different from information that people write solutions about.

10:18 AM - 24 Sep 2018

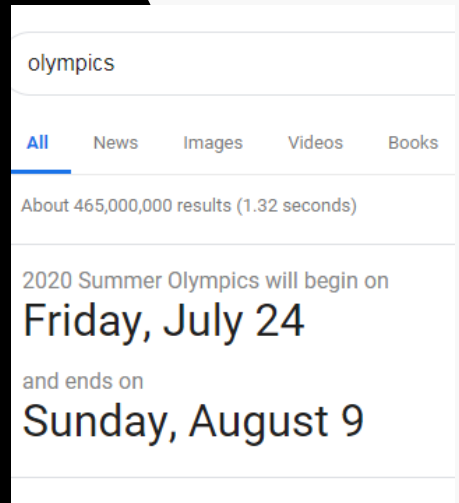
Danny Sullivan ✓ @dannysullivan · 13h

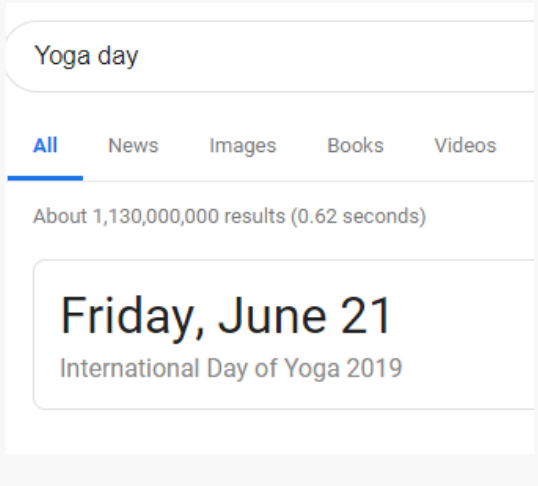
Last few months, Google has been using neural matching, --AI method to better connect words to concepts. Super synonyms, in a way, and impacting 30% of queries. Don't know what "soap opera effect" is to search for it? We can better figure it out.

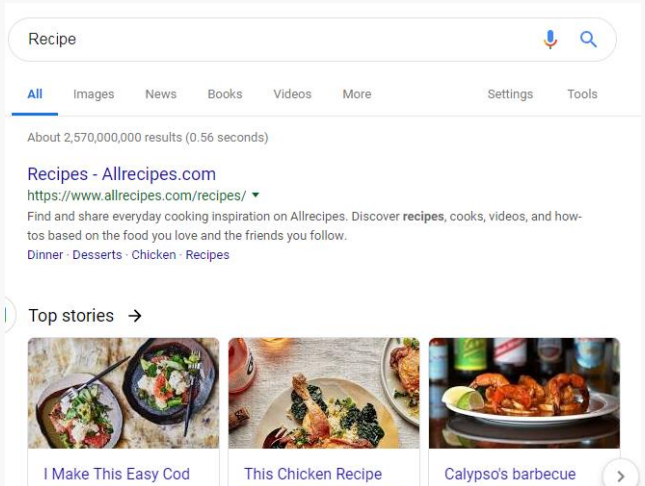


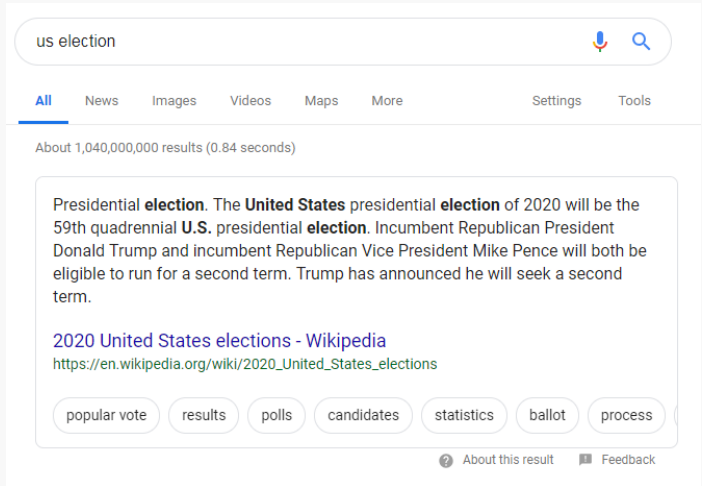
POSSIBLE EXPECTANCY MATRIX

Search Term	Past Result	Future Result	Present Result
Score			Expected to See the Score of Game in the Resident country
Election		We are in 2019 , We may not be more interested in election of the Past, but more current or Future election.	
Recipe	Valid to include such results , as it won't change much.		Also any trending food's recipe
Event		All possible future events of the country.	Expected to see all possible Event of the geo graphical location or country .
International Day/ Event / Sport		Showing the Future most current date of the event	



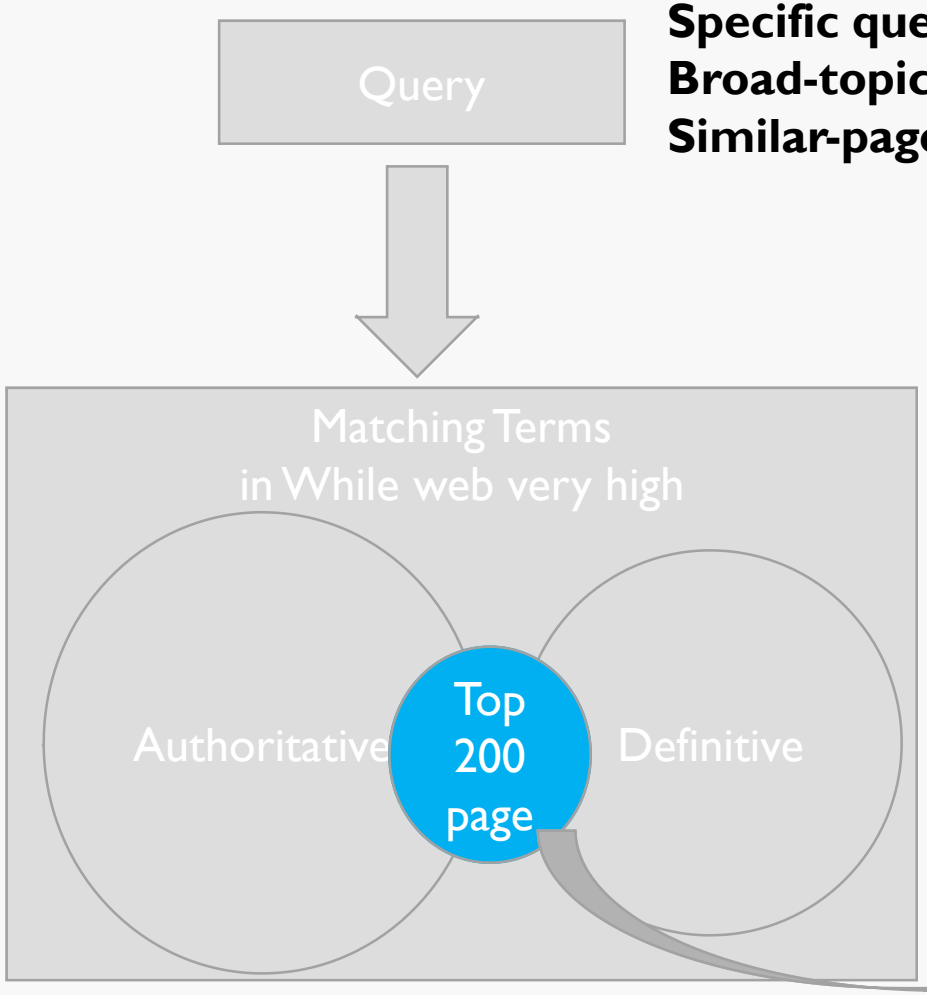






SCENARIO DESIGN : SEARCH RESULT

Specific queries. E.g., "Does Netscape support the JDK 1.1 code-signing API?"
Broad-topic queries. E.g., "Find information about the Java programming language."
Similar-page queries. E.g., "Find pages 'similar' to java.sun.com."



Increase Frequency by ONE for each page that is Present in each other. This would result in page with most referred among the top 200.

This is very basic view, since high frequency doesn't guarantee that you will have great content.

APPENDIX

https://vannizhang.github.io/wonder/#

https://www.youtube.com/embed/IdjJgd6gezM

https://arxiv.org/pdf/1711.08611.pdf
https://www.cs.cornell.edu/home/kleinber/auth.pdf
http://www.bayardo.org/ps/vldb2009.pdf
https://blog.monitorbacklinks.com/seo/recover-google-panda-penalty/
https://www.forbes.com/sites/lutzfinger/2014/09/02/recommendation-engines-the-reason-why-we-love-big-data/#589fafba1077

https://www.google.com/search/howsearchworks/algorithms/

https://search.googleblog.com/2011/11/giving-you-fresher-more-recent-search.html

https://googleblog.blogspot.com/2011/11/giving-you-fresher-more-recent-search.html

https://googleblog.blogspot.com/2010/06/our-new-search-index-caffeine.html