
Search 101

August 23, 2019,
Maish Nichani (maish@pebbleroad.com)



Agenda

- About me
 - About search
 - 4 search stories
 - What is enterprise search
 - Elements of search experience
 - Search strategy
 - Search projects
-

What does enterprise search
mean to you?

“Just make it work like Google”

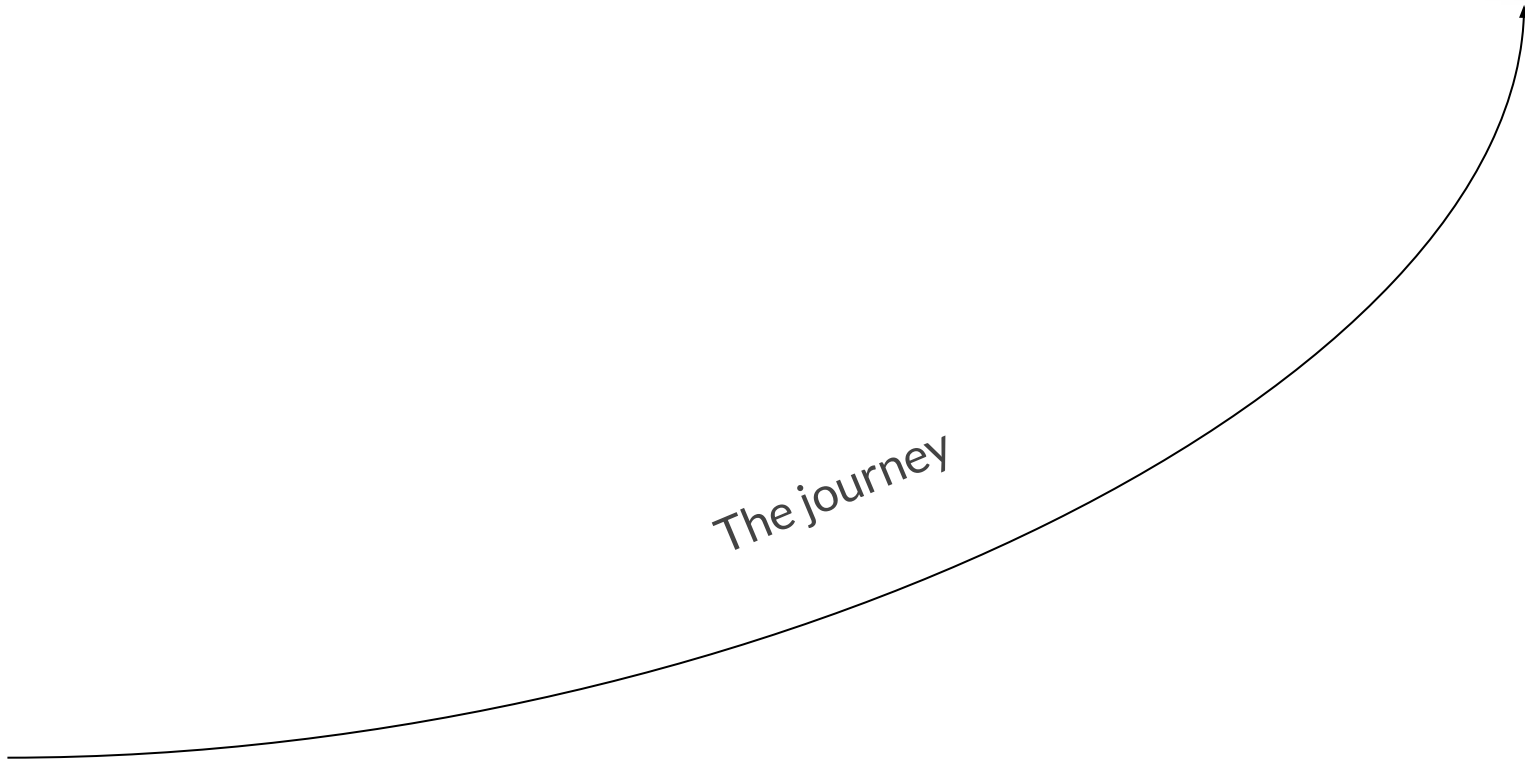




Want search to work like magic



But starts like a black box not many understand



Enterprise search

Enterprise search leverages **powerful stack of technologies** to tame **large volumes of heterogeneous information** in order to offer users a **fruitful search experience**.

Elements of search experience



Search goals



User



Interface



Technology



Context



Content



Creator



Governance

User

We use search to become<one of these>for a while



Assistant



Explorer



Analyst



Executive

Assistant



How many pounds in a kg?

Do I need a visa to enter Myanmar?

What does this transaction code mean?

Explorer



What do I advertise for Java developers?

Who sells good wooden furniture?

What are my employment benefits?

Analyst



Which is the best phone to buy for dad?

Does that company really sell 'fair-trade' coffee?

How does the VW Golf compare with other cars in its class?

Executive

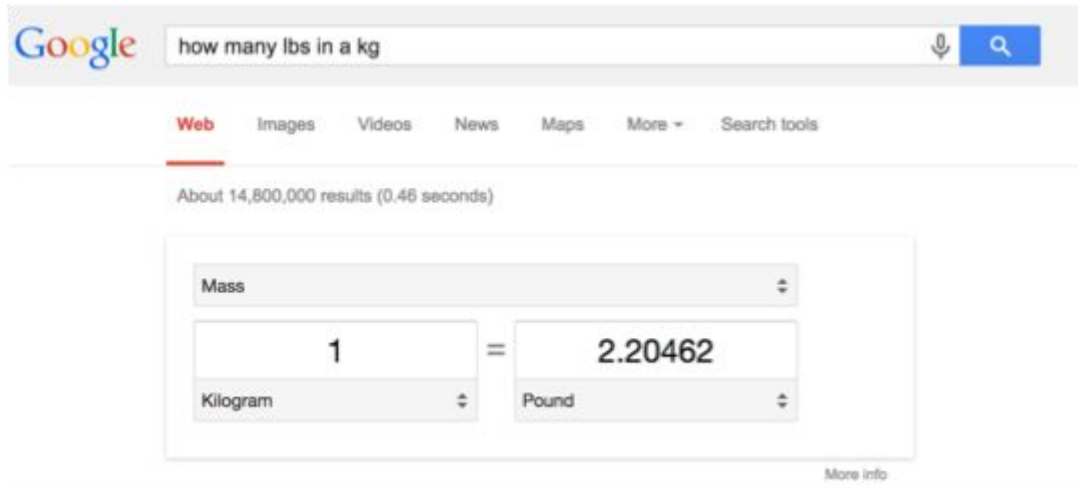


How has sales fared since the apology for cheat software in our cars?

How many cases were opened this year and how many of them closed?

Which areas in Singapore have the highest office rental yield?

How many pounds in a kg?



A screenshot of a Google search result for the query "how many lbs in a kg". The search bar at the top contains the text "how many lbs in a kg" and a blue search button. Below the search bar, the "Web" tab is selected, and navigation links for "Images", "Videos", "News", "Maps", "More", and "Search tools" are visible. The search results indicate "About 14,800,000 results (0.46 seconds)". A unit converter widget is displayed, showing a conversion from Kilogram to Pound. The widget has a dropdown menu set to "Mass" and an equals sign between two input fields. The left field contains the value "1" and is labeled "Kilogram" below it. The right field contains the value "2.20462" and is labeled "Pound" below it. A "More info" link is located at the bottom right of the widget.

Google how many lbs in a kg

Web Images Videos News Maps More Search tools

About 14,800,000 results (0.46 seconds)

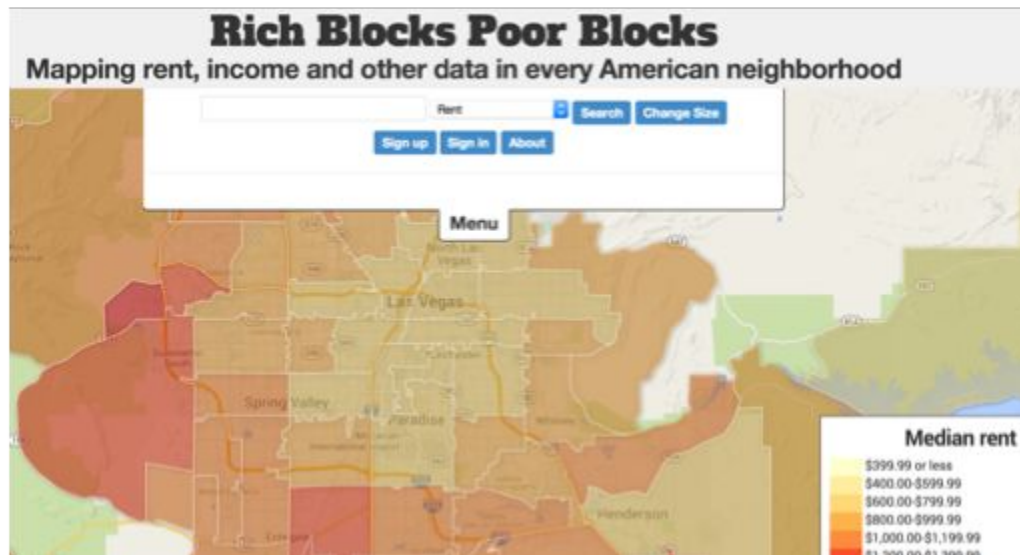
Mass

1 = 2.20462

Kilogram Pound

More info

Which areas will give highest rental yield?



Which phone to buy for dad?

1-24 of 1,353 results for Cell Phones & Accessories: Cell Phones **No-Contract Cell Phones** Sort by: New and Popular

Network
☐ 3G (471)
☐ GSM (373)
☐ 4G (181)
☐ Quad Band (113)
☐ Dual Band (33)
☐ Tri Band (3)

Internal Storage
☐ 4 GB (132)
☐ 8 GB (96)
☐ 16 GB (116)
☐ 32 GB (97)
☐ 64 GB (15)

Feature Keywords
☐ QWERTY (187)
☐ Straight Talk (18)

Avg. Customer Review
★★★★★ & Up (396)
★★★★☆ & Up (374)
★★★☆☆ & Up (367)
★★☆☆☆ & Up (357)

International Shipping
☐ AmazonGlobal Eligible

Condition
New (1,111)
Used (218)
Refurbished (302)

Price
Under \$10 (16)
\$10 to \$24 (193)

Nokia Lumia 520 GoPhone (AT&T)
\$249.99 **\$39.99** iPhone
Get it by **Monday, Nov 17**
FREE Shipping
More Buying Choices
\$39.99 new (83 offers)
\$32.99 used (21 offers)
★★★★☆ (3) > 12,028
[Get Deal Now](#) in Prepaid Phones
[See newer version](#)

HTC Desire 816 - Prepaid Phone (Virgin mobile)
\$249.99 **\$252.75** iPhone
Get it by **Monday, Nov 17**
FREE Shipping and 1 more promotion >>
More Buying Choices
\$279.99 new (3 offers)
★★★★☆ (3) > 4,000

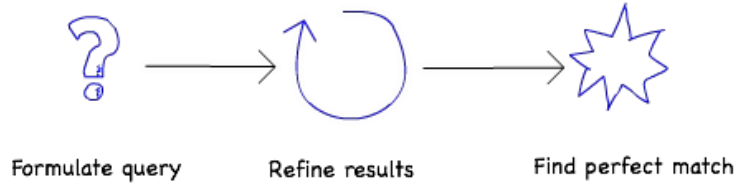
LG Volt White - Prepaid Phone (Virgin Mobile)
\$199.99 **\$99.99** iPhone
Get it by **Monday, Nov 17**
FREE Shipping and 1 more promotion >>
More Buying Choices
\$99.99 new (18 offers)
\$84.99 used (1 offer)
★★★★☆ (3) > 228

LG Volt - Prepaid Phone (Boost Mobile)
\$199.99 **\$79.65** iPhone
Get it by **Monday, Nov 17**
FREE Shipping and 1 more promotion >>
More Buying Choices
\$89.99 new (22 offers)
\$79.65 used (1 offer)

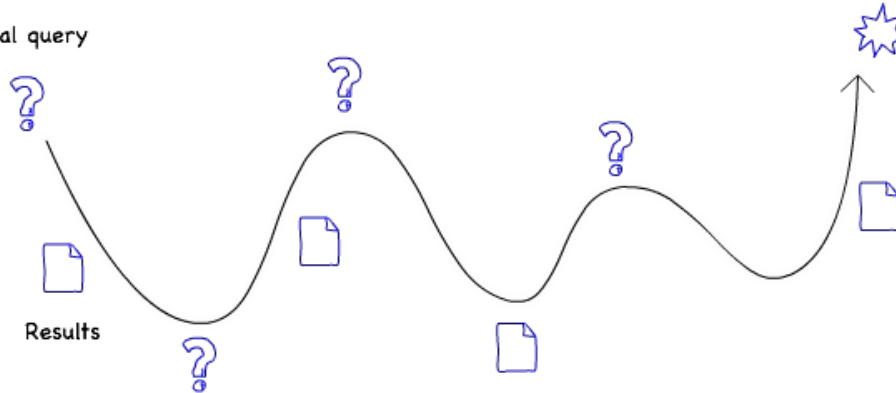
Samsung Galaxy S II Triband (Boost Mobile)
\$249.99 **\$199.99** iPhone
Get it by **Monday, Nov 17**
FREE Shipping
More Buying Choices
\$199.99 new (22 offers)
\$179.99 used (1 offer)

AT&T Nokia Lumia 635 - No Contract GoPhone
\$199.99 **\$79.99** iPhone
Get it by **Monday, Nov 17**
FREE Shipping
More Buying Choices
\$79.99 new (22 offers)
\$69.99 used (1 offer)

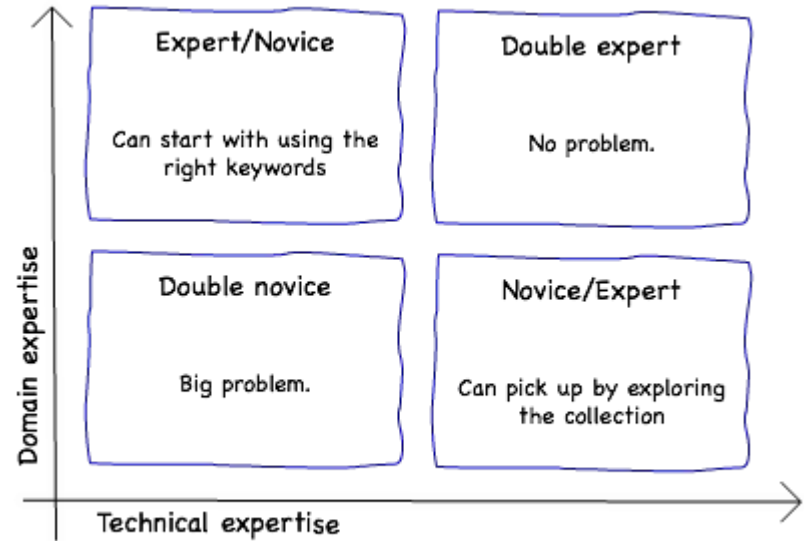
Infoseeking behaviours



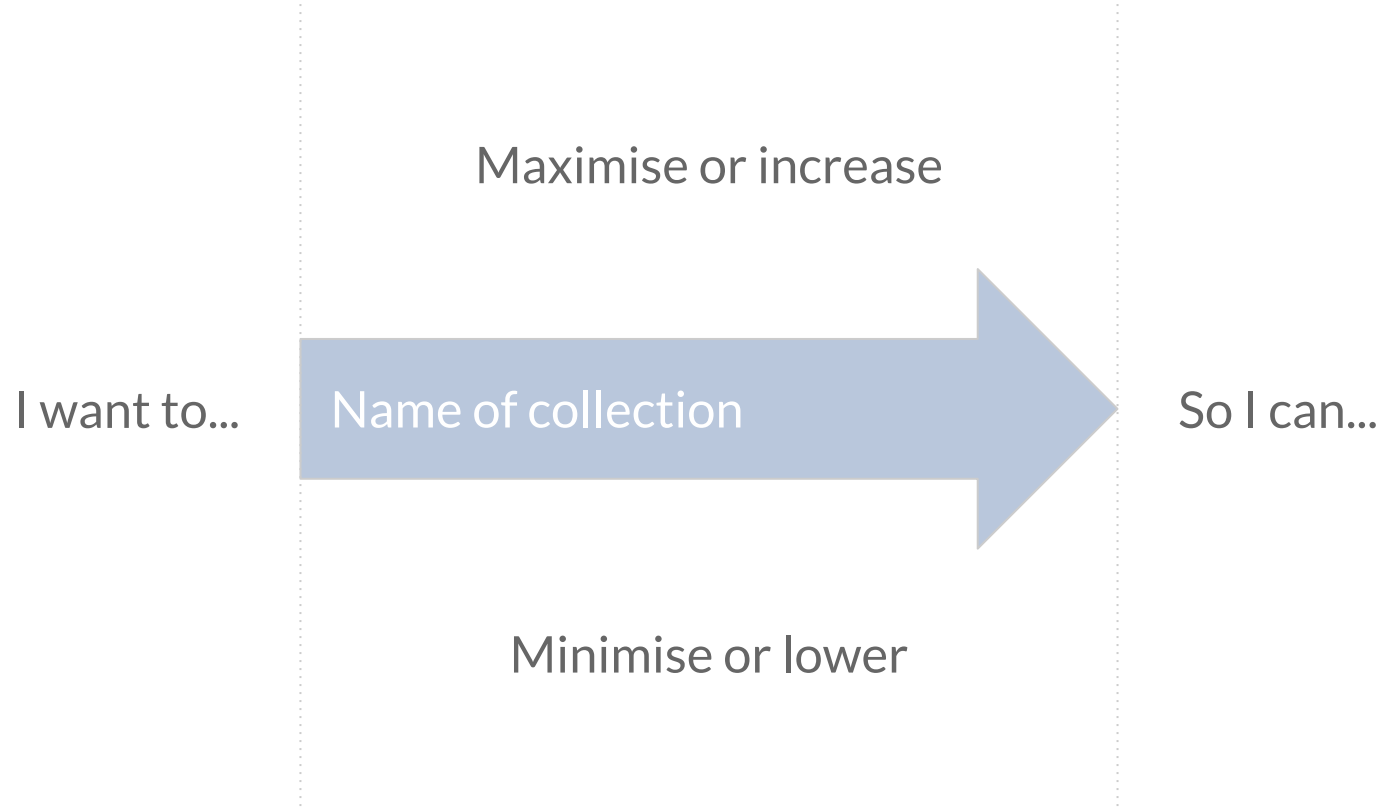
Initial query



Expertise



Persona



Jack

Find the right movie quickly
Watch in good quality resolution

I want to watch a
comedy movie
after a hard day's
work

Netflix

No buffering pauses
No other distractions

Unwind and relax

Content

Content

Google chocolate chip cookie recipe

Web Videos Images Shopping News More Search tools

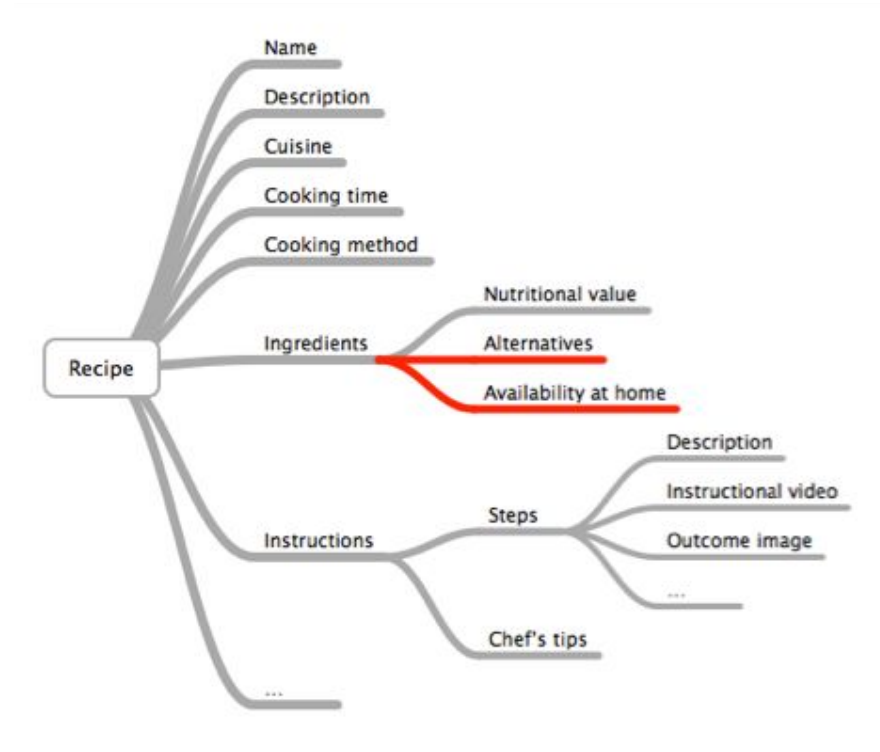
Ingredients Any cook time Any calories Any time All results

Ingredients	Yes	No
Chocolate chips	☐	☐
Baking soda	☐	☐
Vanilla extract	☐	☐
Baking powder	☐	☐
Sugar	☐	☐
Flour	☐	☐
Eggs	☐	☐
Butter	☐	☐

Chocolate Chip Cookies Recipe | Martha ...
<https://www.marthastewart.com/.../soft-and-chewy-...> Martha Stewart
Rating: 3.8 - 281 reviews
chip cookies are a go-to favorite dessert, sweet snack, or
pl treat. With our easy, three-step recipe, you'll have soft and
chocolate chip ...
Chocolate Chip Cookies - Soft-Baked Chocolate Chip ... - Chocolate

Original NESTLÉ® TOLL HOUSE® Chocolate Chip Cookies ...
<https://www.verybestbaking.com/recipes/.../original-nestle-t...>
55 mins
Original NESTLÉ® TOLL HOUSE® Chocolate Chip Cookies This

Modelling for relevance



- 1. No metadata, no search**
- 2. Bad metadata is worse than no metadata**



Agnes Molnar

ATM collection

#	Location	Address
1	Thomson Road	213 Thomson Road Singapore 432123

ATM collection (optimised)

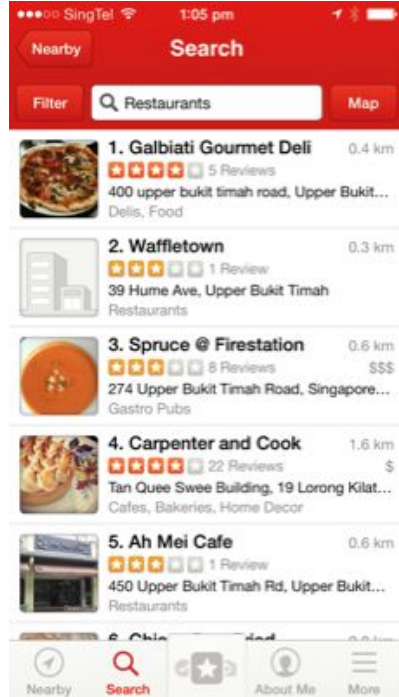
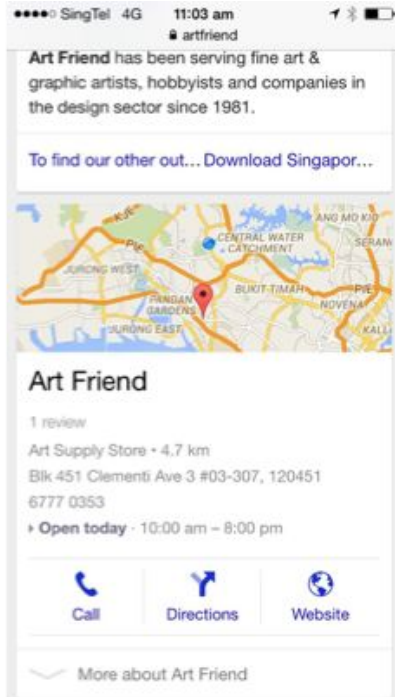
#	Location	Address	System status
1	Thomson Road	213 Thomson Road Singapore 432123	Out of order

Content type	Employment agencies
Description	A list of employment agencies in Singapore.
Core search task	To lookup employment agencies by name, region and ranking.

Metadata	Description	Values	Source	How it can help
Name				
Region	The region that the agency falls under.	North South East West Central		With only the addresses the workers don't have much to select agencies. With regions they can narrow down on the agencies near their residence or dorms.
Address				
Location zip code				
Contact phone number				

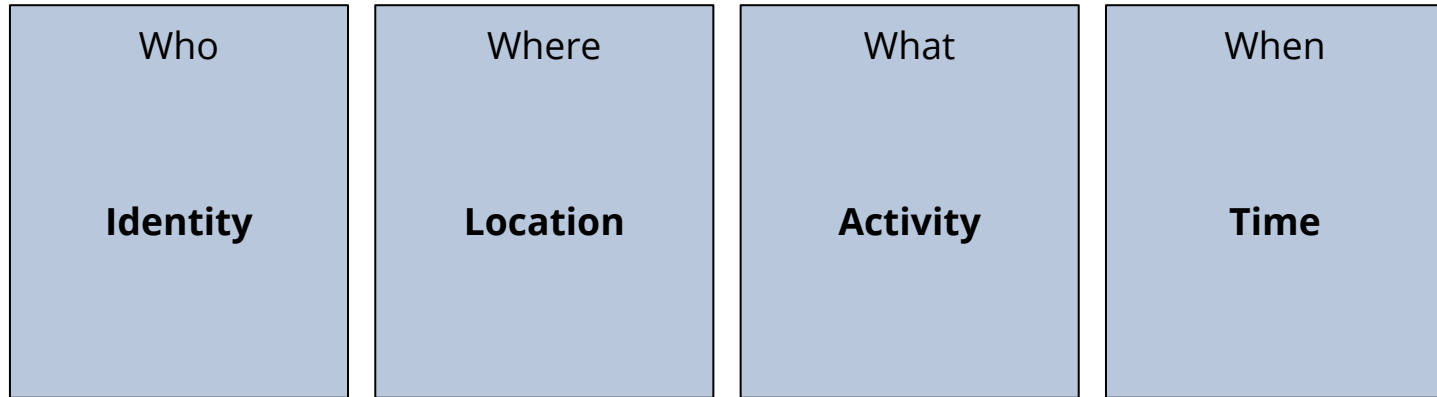
Context

Context



Searching from Bukit Timah

Context



Towards a Better Understanding of Context and Context-Awareness,
<https://smartech.gatech.edu/bitstream/handle/1853/3389/99-22.pdf>

Persona

Scenario

Device

Environment

Time

Activity

Individual

Location

Social

Jack, the travelling salesman

Device
Environment
Time
Activity
Individual
Location
Social

Scenario

When Jack is in office and searching for a topic from the projects page, results from projects similar to the one he is working on right now are ranked higher.

If Jack is new to the company, in addition to the search results, results from the How-to section are ranked higher.

Interface

Interface

The beautiful box



Enter keywords

Microphone icon, Heart icon, Magnifying glass icon

Interface

Input

Features that allow the searcher to express what they are looking for (e.g., search bar)

Control

Features that help searchers to modify, refine, restrict or expand their Input (e.g., filters)

Informational

Features that provide results or information about results (e.g., snippets)

Personalisable

Features that relate specifically to searchers and their past interactions (e.g., bookmarks)

1-16 of over 100,000 results for "mobile phones"

Sort by: Featured ▾

Department

Cell Phones & Accessories

Carrier Cell Phones

Cell Phone External Flashes & Selfie
Lights

Books

Computers & Technology

Mobile & Wireless Computing

▾ See All 31 Departments

Avg. Customer Review

★★★★★ & Up

★★★★☆ & Up

★★★☆☆ & Up

★★☆☆☆ & Up

Brand

☐ Mobile Phone☐ Mobile Phones☐ Samsung☐ Nokia☐ HTC☐ Jabra☐ CyberTech☐ KONOVA☐ Reiko☐ Graphics and More☐ Choseal☐ GooIT

SPONSORED BY LG ELECTRONICS

Sleek, new
technology with
touchless commands[Shop now ▸](#)LG G8 ThinQ with Alexa Hands-Free –
Unlocked SMARTPHONE – 128 GB – Auror...

★★★★☆ 26

✓prime

[Ad feedback](#) 🗨

Sponsored ⓘ

NUU Mobile G3 Unlocked Cell Phone 64GB + 4GB RAM - 5.7" Android
Smartphone - Sapphire Blue

★★★★☆ ▾ 112

\$139⁹⁹

Eligible for Shipping to Singapore



Sponsored ⓘ

NUU Mobile G3 Plus Unlocked Cell Phone 64GB + 4GB RAM 4000 mAh



twitter stock



Web Images Videos News Maps **Stock**

Settings ▾

Twitter Inc

TWTR as of Aug 21 4:00 PM ET | Markets are closed

42.80 ▲ **0.51** | **1.21%**

More at NASDAQ | Data by Xignite



twitter stock



Finance



News



Images



Videos



More



Settings



Tools

About 7,430,000,000 results (0.61 seconds)

Twitter Inc

NYSE: TWTR

42.80 USD **+0.51 (1.21%)** ↑

Closed: 21 Aug, 7:59 pm GMT-4 · Disclaimer

After hours 42.88 **+0.080 (0.19%)**

1 day

5 days

1 month

6 months

YTD

1 year

5 years

Max



Open	42.83	Div yield	-
High	43.28	Prev close	42.29
Low	42.36	52-wk high	43.48
Mkt cap	33.09B	52-wk low	26.19
P/E ratio	14.18		

what is the capital of india

Showing results for what is the capital of India



New Delhi

India

Source: [DBpedia](#)

Image of Asia: A refugee in New Delhi

4 September 2015

In this photo by Tsering Topgyal, a Rohingya refugee sews inside a refugee agency, around 9,000 Rohingya refugees are registered in

How does the search engine know that the capital of India is Delhi?

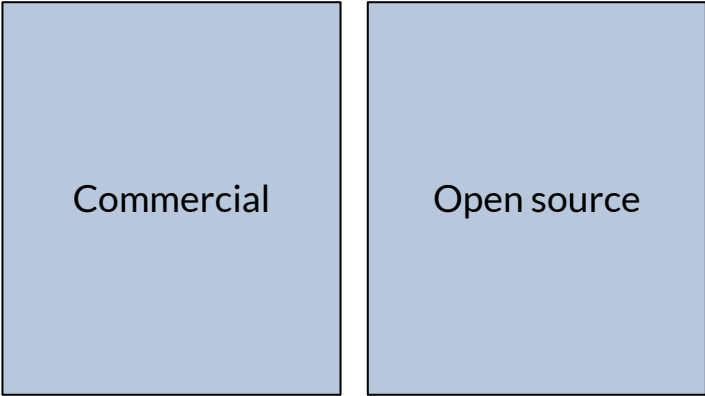
Autocomplete	Autocomplete will attempt to complete the search term just as the searcher starts typing.
Autosuggest	Autosuggest will suggest resources found in the collection to the user while the user is still typing their search terms. It helps speed up retrieval.
Best bets	Best bets will elevate a document to very top of the list, ignoring any ranking algorithms. Can be useful for promoting certain pages on a website.
Context	Context plays a large part in improving relevancy and provides a personalised experience. Includes location, time, user profiles, weather, personal preferences, and other attributes to influence the search ranking algorithm for a custom result set tailored for the user.
Did you mean	A subset of spell checking, 'did you mean' can provide alternate terms to search for that closely matches the original term. Did you mean provides helpful clues for finding resources.
Faceted navigation	Facets act like filters to narrow down the results to find specific resources. This approach supports the exploration and discovery of information.
Field collapsing	When multiple documents are very similar in nature (eg, same movie, different format), consider collapsing the documents by a particular field, such as their title, to return greater variety in the search results.
Hit highlighting	Highlights query terms found in the resources. Helps users to quickly scan results to determine which ones are worth investigating further. Very useful for large documents.
Instant answers	Giving answers in the search results page or in autosuggest. Saves the user time from having to visit pages to get direct answers. This is very useful in mobile contexts.

Instant results	Similar to autosuggest, instant results will automatically update the search results page with new results when a character is typed into the search field. Saves time from having to type the full word or query.
Onboarding	Guides users through the search experience with smart tooltips. Helps novice users become more savvy with search.
Recommendations	Provide additional, similar documents based off of a certain criteria that can be customised.
Results bookmarking	Save a search snippet as as bookmark for easy retrieval at a later date.
Results grouping	Groups resources by specific categories in the search results. This helps to view different types of resources returned for the same term.
Results ranking	Providing options to rank the results by a specific criteria. Most common is by relevance score but others include dates, prices, ratings.
Rich snippets	Search snippets that present a rich set of relevant content to the user based. Often based on user persona (assistant, explorer, analyst, executive).
Search history	Automatically save the trail of searches made for easy retrieval at a later date.
Spell checking	Provides alternate spelling of words in a query to correct spelling mistakes. Helps to improve usability of the search experience. Very useful in mobile contexts.
Visualisations	View a bird's eye view of your collection in various chart types to see a bigger picture on the collection of information. Useful for the analyst user persona.

Technology

Technology

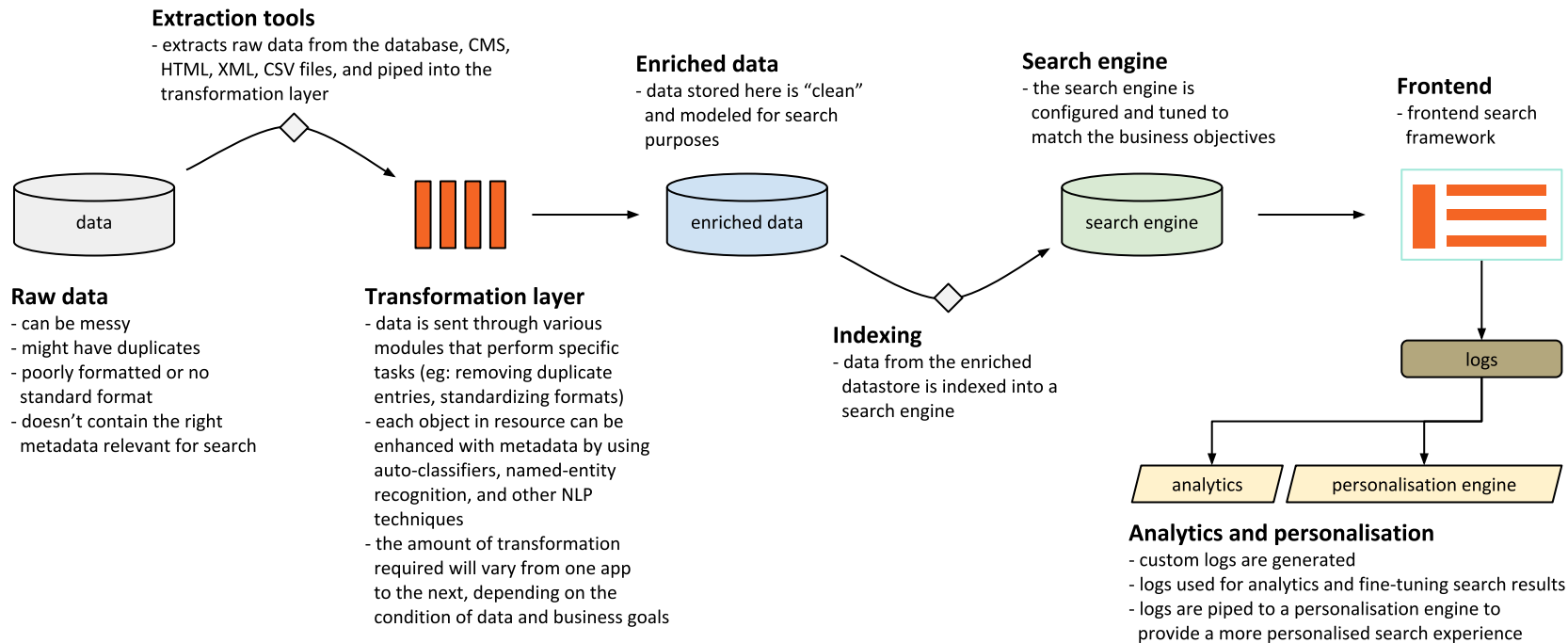
Dates
Duplicates
Entity extraction
Faceting
Multi-lingual
People search
Query management
Ranking
Sentiment analysis
Spell checking
Stemming and lemmatization
Stop words
Summarisation
Tokenization
...



Commercial

Open source

Technology



Technology

1. Strength
2. Scalability
3. Satisfaction - ease of use

You may find that open source search engines are equal to, if not better than, commercial ones.



All together

Search goal: Apply for paternity leave.

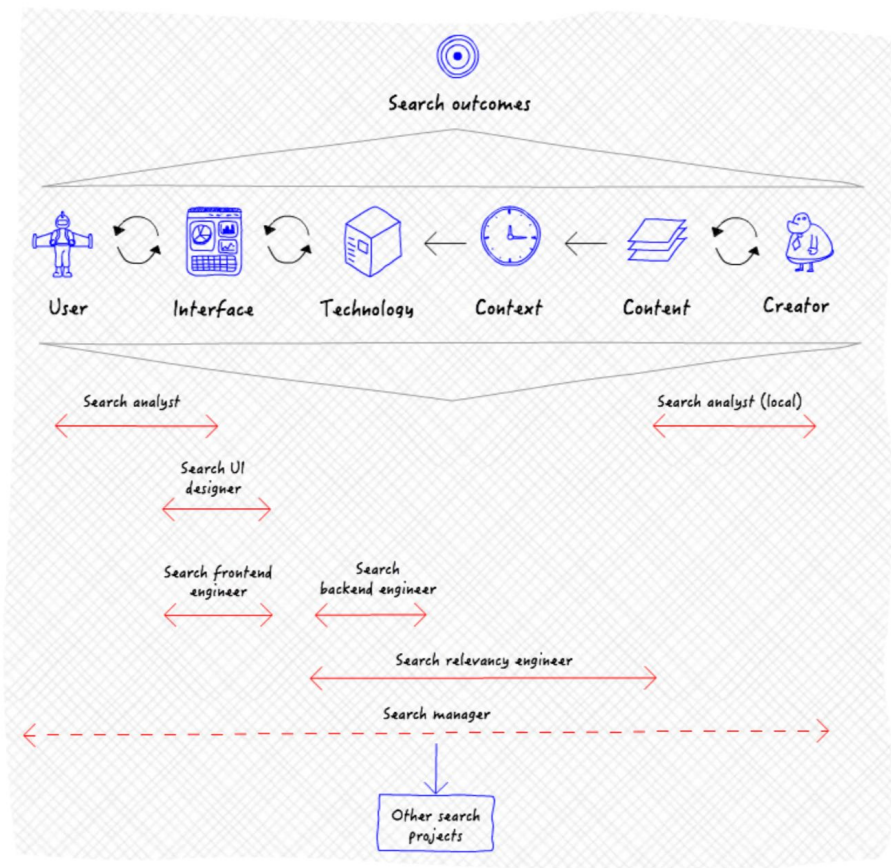
Scenario	Example
1. The result snippet shows the title and content around the keywords.	HR@YourCompany Leave Policy http://internal.yourcompany.com/policies/leave/ You can find all policies related to leave in one place. How cool is that?
2. The result snippet shows the title and matching subhead on paternity leave.	HR@YourCompany Leave Policy http://internal.yourcompany.com/policies/leave/ Not only can you find all policies in one place, but you can also navigate to the policy you want right here. How cool is that? Annual leave Sick leave Maternity leave Paternity leave Childcare leave ...
3. The result snippet tells how many days of paternity leave the user is eligible for.	Hi John Malkovich, based on your records You are eligible for 12 days of paid paternity leave. Apply for paternity leave Or find out about the conditions of leave.

Governance

Governance

"The impact of search on business performance depends more on the level of investment in a skilled team of people to support search than it does on the level of investment in search technology."

Martin White, author of *Enterprise Search*

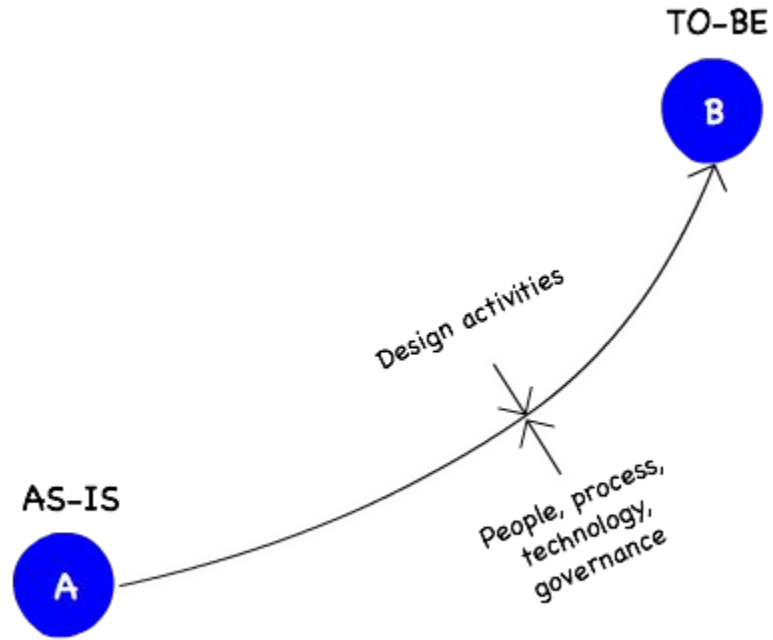


How is search performance measured?

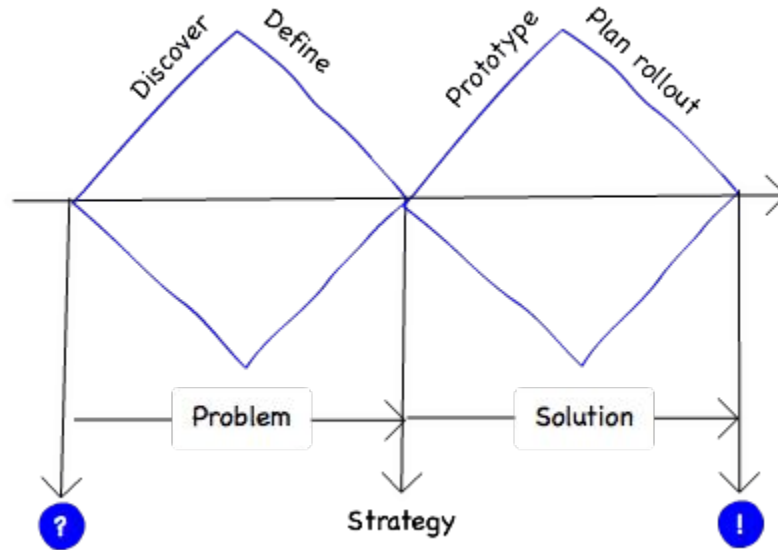
- How many users are using the search box?
- What are the top xx places users begin their searches?
- What are the top xx search queries?
- What are the top xx pages reached through search?
- What are the queries with 0 results?
- What are the queries with 0 click-throughs on the SERP?
- How many times was search used immediately after the first SERP?
- What are the queries that result in more than 3 SERPs?
- How many users used facets or filters?
- How many users use “best bets”?
- ...

Planning a enterprise search project

The plan



The process



The activities

Phase	Activities
Discover the opportunity	Review target resource collection Interview target users Review search interfaces Analyse search logs Review search technologies Review search governance
Define the desired experience	Identify dominant search personas Identify core search tasks Identify content models for relevance Identify search interface features Identify core technology requirements Identify core measurement metrics Define search governance requirements
Design a search prototype	Prepare resource collection for prototype Design search interfaces Deploy search technology for prototype Design and test search prototype
Develop the new search service	Prepare actual resource collection Implement search technologies Develop search interfaces Configure the search system Test system with actual users
Govern the new experience	Analyse search clickstream logs Analyse site search logs Conduct search usability tests

Thank you

