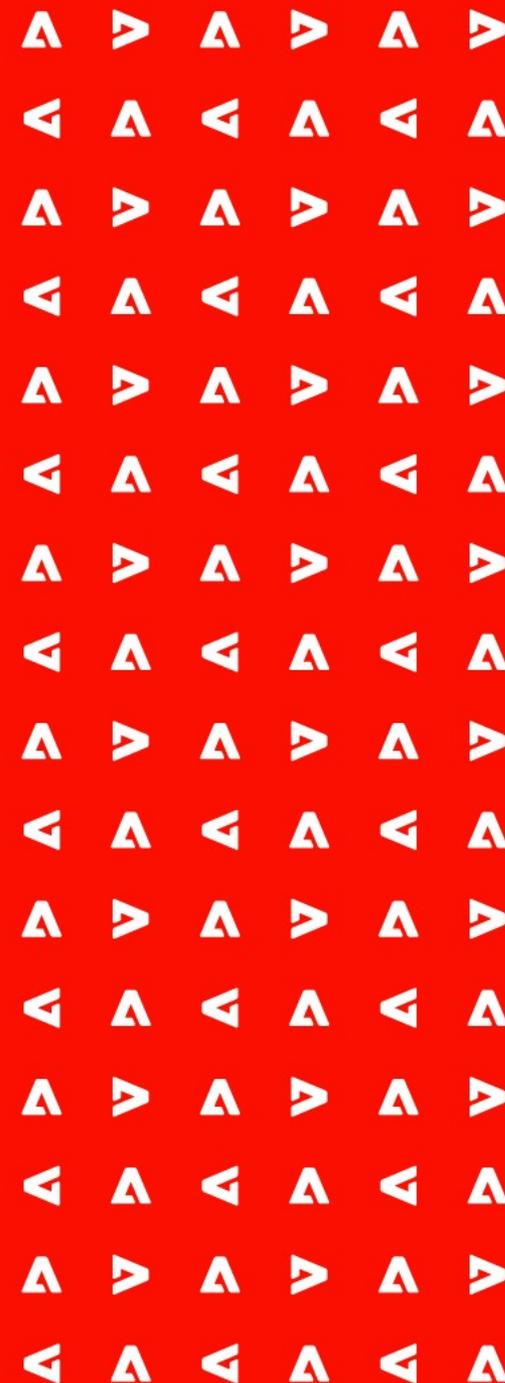




White Roses, Red Backgrounds: Bringing Structured Representations to Search

Tracy King, Senior Principal Scientist
Adobe Sensei & Search

2021-12-15 LACompLing2021



white rose red background



white rose red background



snacks without nuts



Sponsored ⓘ

KIND Bars, Dark Chocolate Nuts & Sea Salt, Gluten Free, 1.4 Ounce Bars, 24 Count

24 Count (Pack of 1)

★★★★★ ∨ 2,632

\$28⁴⁴ (\$1.19/Count)

\$27.02 with Subscribe & Save discount

SNAP EBT eligible

✓prime FREE Delivery **Sat, Dec 11**



Macadamia Nut Butter| 100% Pure Raw and Without Any Additives or Sugar| Made from 100% Macadamia Nuts| Perfect Keto Diet Snack - 12 Oz

12 Ounce (Pack of 1)

★★★★☆ ∨ 130

\$29⁴⁹ (\$2.46/Ounce)

✓prime FREE Delivery **Sun, Dec 12**

Only 6 left in stock - order soon.

More Buying Choices

\$24.45 (2 new offers)



Without A Trace Top 9 Allergen-Friendly, Nutrient Dense, Plant Based Power Bites - Nut Free Snacks, Gluten Free And Snack Size. Convenient &...

★★★★☆ ∨ 52

\$26⁹⁰ (\$2.13/Ounce)

✓prime FREE Delivery **Sun, Dec 12**



Deluxe Roasted Salted Mixed Nuts in Sea Salt by Its Delish – Bulk 2 lbs - Premium Roasted Nut Mix Without Peanuts – Delicious Snacking Nuts –...

★★★★☆ ∨ 52

\$24⁹⁹ (\$24.99/Count)

Get it **Fri, Dec 17 - Tue, Dec 21**

FREE Shipping

snacks without nuts vs. nut free snacks



Sponsor

KIND Peanut Free & Tree Nut Free Snack
Sea Sack - Back to School Snack Variety
24 Count Pack or College Care Package |
24 Count Popcorn, Chips, Cookies, & Sweets - ...

★★★★★ 42 Piece Set

\$284.99
\$27.08/Count
★★★★★ ~ 20

Save 5%

\$34.95 (\$34.95/Count) \$36.95

Save more with Subscribe & Save
Lowest price in 30 days

✓prime FREE Delivery Sun, Dec 12



Made Good Healthy Snacks Variety
Pack, 40 Ct - Organic Assortment of
Granola Bars, Granola Mini Snack
Packs, Crispy Squares; Individually...

7 Piece Assortment

★★★★★ ~ 2,502

\$29.99 (\$29.99/Count)

\$28.49 with Subscribe & Save discount

✓prime FREE Delivery Sat, Dec 11

Climate Pledge Friendly

See 1 certification



Allergy Friendly GLUTEN FREE and
VEGAN (15Ct): DAIRY, WHEAT,
PEANUT, TREE NUT, SOY, and EGG
FREE) Healthy Snacks Care Package...

15 Piece Set

★★★★★ ~ 6

\$39.95 (\$2.66/Count)

Save more with Subscribe & Save

✓prime FREE Delivery Sat, Dec 11



Enjoy Life Foods Seed and Fruit Mix,
Peanut Free Trail Mix, Mountain
Mambo, 39.12 Ounce (Pack of 24)
1.63 Ounce (Pack of 24)

★★★★★ ~ 1,099

\$32.98 (\$0.84/Ounce)

\$31.33 with Subscribe & Save discount

40% off your first subscription order
SNAP EBT eligible

✓prime FREE Delivery Sat, Dec 11

More Buying Choices

\$32.50 (6 new offers)

dresses over \$100



Marina Womens Gown Blue Size 6
Cold Sholder Cape Over Lay Lace...
New (Other)

\$32.49

List price: ~~\$100.00~~ **68% off**

or Best Offer

Free shipping

Free returns

Sponsored



\$35.19

Was: ~~\$39.99~~ **12% off**

or Best Offer

+\$6.95 shipping

Sponsored



J Crew Sz 0 All Over Eyelet Belted
Dress Blue 100% Cotton Fit & Flare...
Pre-Owned

\$9.99

0 bids · 4d 15h

or Best Offer

+\$8.79 shipping

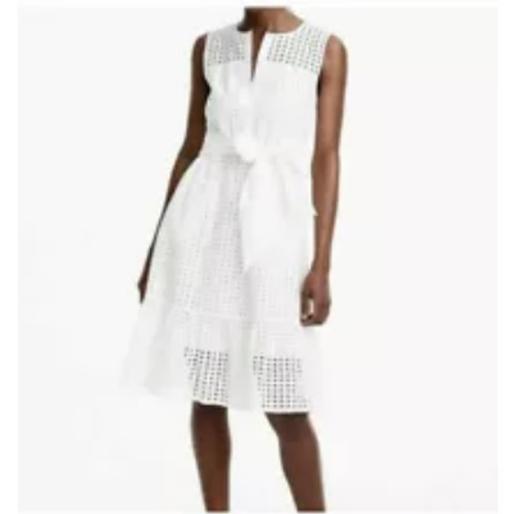


Nine West Fully Lined Dress Size 8.
EUC. absolutely Adorable. MSRP...
Pre-Owned

\$54.97

or Best Offer

Free shipping



J.Crew All Over Eyelet Lace White
Tie Dress J1608 100% Cotton Size ...
Brand New

dresses over \$100



Alex Evenings

Women's Plus Size Long A-Line
Rosette Dress with Short Sleeves

★★★★★ 650

\$249⁰⁰



Prime Try Before You Buy



Tahari ASL

Women's Long Sleeve Sequin Surplus
Wrap Dress

★★★★★ 6

Prime Try Before You Buy



Norma Kamali

Women's Long Sleeve Crew Fishtail
Dress to Midcalf

★★★★★ 1

\$158⁸⁶



Alex Evenings

Women's Plus Size Long Dress with
Mandarin Neckline Jacket

★★★★★ 490

\$185⁰⁰ ~~\$205.00~~



Prime Try Before You Buy

Mar
Col
New

\$32

List
or B
Free
Free
Sp

Talk outline

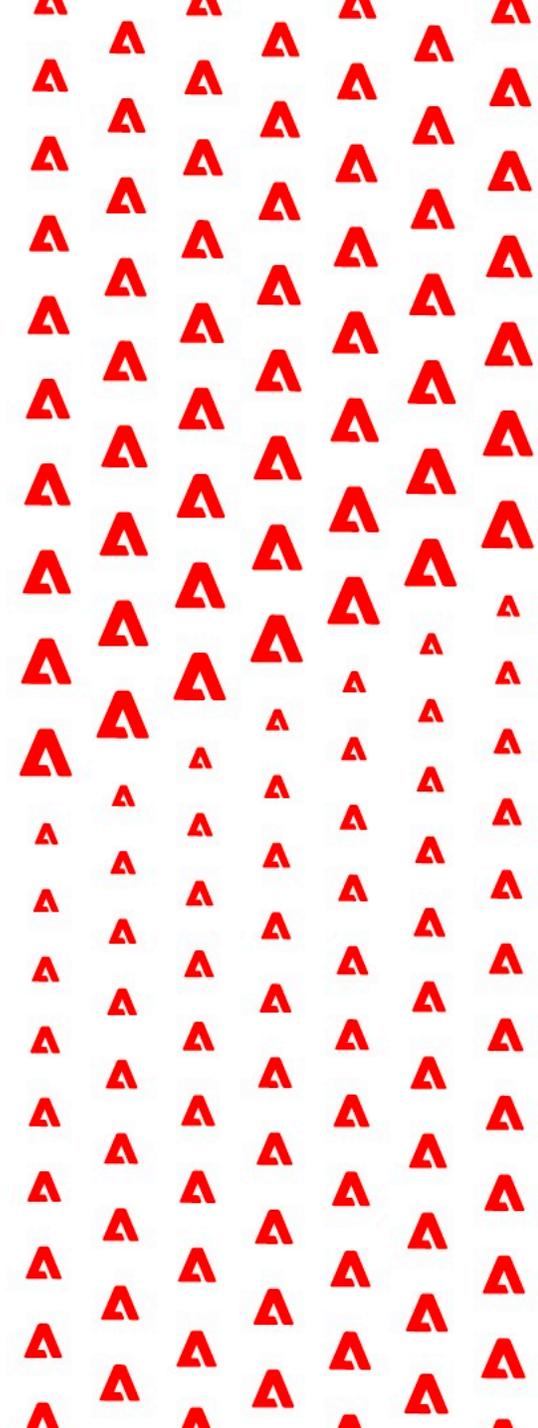
- How search works
- Where this fails us
- Towards a solution



 <p>Alex Evenings Women's Plus Size Long A-Line Rosette Dress with Short Sleeves ★★★★☆ ~ 650 \$249⁰⁰ prime Prime Try Before You Buy</p>	 <p>Tahari ASL Women's Long Sleeve Sequin Surplus Wrap Dress ★★★★☆ ~ 6 Prime Try Before You Buy</p>	 <p>Norma Kamali Women's Long Sleeve Crew Fishtail Dress to Midcalf ★★★★☆ ~ 1 \$158⁸⁶ prime</p>	 <p>Alex Evenings Women's Plus Size Long Dress with Mandarin Neckline Jacket ★★★★☆ ~ 490 \$185⁰⁰ \$205.⁰⁰ prime Prime Try Before You Buy</p>
--	---	--	--

 <p>Peanut Free & Tree Nut Free Snack Sack - Back to School Snack Variety Pack or College Care Package Popcorn, Chips, Cookies, & Sweets ... 42 Piece Set ★★★★☆ ~ 20 Save 5% \$34⁹⁵ (\$34.95/Count) \$36.⁹⁵ Save more with Subscribe & Save Lowest price in 30 days prime FREE Delivery Sun, Dec 12</p>	 <p>Made Good Healthy Snacks Variety Pack, 40 Ct - Organic Assortment of Granola Bars, Granola Mini Snack Packs, Crispy Squares; Individually... 7 Piece Assortment ★★★★☆ ~ 2,502 \$29⁹⁹ (\$29.99/Count) \$28.49 with Subscribe & Save discount prime FREE Delivery Sat, Dec 11 Climate Pledge Friendly See 1 certification</p>	 <p>Allergy Friendly GLUTEN FREE and VEGAN (15C): DAIRY, WHEAT, PEANUT, TREE NUT, SOY, and EGG FREE) Healthy Snacks Care Package... 15 Piece Set ★★★★☆ ~ 6 \$39⁹⁵ (\$2.66/Count) Save more with Subscribe & Save prime FREE Delivery Sat, Dec 11</p>	 <p>Enjoy Life Foods Seed and Fruit Mix, Peanut Free Trail Mix, Mountain Mambo, 39.12 Ounce (Pack of 24) 1.63 Ounce (Pack of 24) ★★★★☆ ~ 1,099 \$32⁹⁸ (\$0.84/Ounce) \$31.33 with Subscribe & Save discount \$20 off your first subscription order SNAP EBT eligible prime FREE Delivery Sat, Dec 11 More Buying Choices \$32.50 (6 new offers)</p>
--	---	--	---

A whirlwind course on how search works



Standard search: Inverted index

A search index is like an index at the back of a book (hence the name).

1. For every “document”, find all the words associated with it
2. Build an index with the words & each document that contains them
3. Break the query into words
4. Find all the documents associated with each word
5. Return the documents that have all the words of the query
6. Rank them

variables, 376, 514
variable binding, 520
verb-initial clauses, 46
verb-initial languages, 417–420
VerbNet, 251, 252, 257–272, 284
verbs, 259, 417, 419, 421, 422,
428–436, 449–451, 457, 459,
515, 519, 522, 550, 554
veridicality, 262, 263, 292
veridicity, 261–263
voice, 319, 322, 340
VOS languages, 419
vowel harmony, 390, 494
VP ellipsis, 10
VSO languages, 419

D1

cats are furry



D2

mice are furry



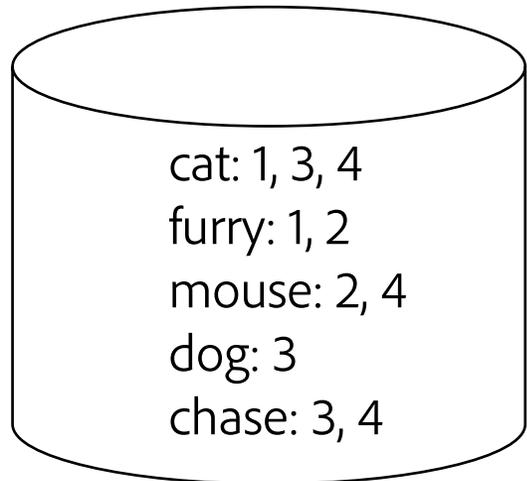
D3

dogs chase cats



D4

cats chase mice



Q: cats



D:cat: 1, 3, 4

Q: cats dogs



D:cat: 1, 3, 4 & D:dog: 3



D: 3

Q: cats chase



D:cat: 1, 3, 4 & D:chase: 3, 4



D: 3, 4

Beyond Just Inverted Indices: Simple text processing

- Lower casing, de-accenting, punctuation stripping, "stop words", lemmatization

Amazon's Choice



IRWIN Hammer, Fiberglass, General Purpose, Claw, 16 oz. (1954889)

★★★★★ ~ 3,425

\$16¹⁷

✓prime FREE One-Day
Get it Tomorrow, Dec 11



CRAFTSMAN Hammer, Fiberglass, 16 oz. (CMHT51398)

★★★★★ ~ 1,929

\$11⁹⁸ ~~\$13.00~~

✓prime FREE One-Day
Get it Tomorrow, Dec 11



STEELHEAD 5-Piece Fiberglass Handle Hammer & Mallet Set, 16oz Ball-Peen Hammer, 32oz Ball-Pein Hammer, 32oz Rubber Mallet, 3lb...

★★★★★ ~ 275

\$49⁹⁹

✓prime FREE Delivery Mon, Dec 13

Best Seller



Mr. Pen- Hammer, 8oz, Small Hammer, Camping Hammer, Claw Hammer, Stubby Hammer, Tack Hammer, Hammers Tools, Small...

★★★★★ ~ 2,330

\$6⁹⁹

✓prime FREE Delivery Sun, Dec 12

hammers

Beyond Just Inverted Indices: Memorizing results

- User behavioral data



The screenshot shows a Google search for 'bbc'. The search bar contains 'bbc' and the search button is visible. Below the search bar, there are navigation tabs for 'All', 'News', 'Videos', 'Images', 'Books', and 'More'. The search results show 'About 3,990,000,000 results (0.66 seconds)'. The first result is 'https://www.bbc.com' with the title 'BBC - Homepage' and a description: 'Breaking news, sport, TV, radio and a whole lot more. The BBC informs, educates and entertains - wherever you are, whatever your age.' Below this, there are two columns of results: 'News' with a snippet 'But UK tests show a booster vaccine dose is key, offering 75 ...' and 'World' with a snippet 'Get the latest BBC World News: international news, features ...'. A 'More results from bbc.com »' link is also present. On the right side, there is a 'See results about' sidebar with three entries: 'BBC Broadcasting company', 'BBC News TV network', and 'BBC Company'. At the bottom, there is a section titled 'Latest from bbc.com' with three image thumbnails: a person's face, a grey t-shirt with 'BRISTOL' written on it, and a group of people holding a sign that says 'KEEP YOUR HANDS OFF OUR BODIES'.

Beyond Just Inverted Indices: Meta-data & structured data

- Documents contain information beyond text (and images)
 - Products: Price, brand, size
 - Publications: Author, year, publisher

dresses over \$100

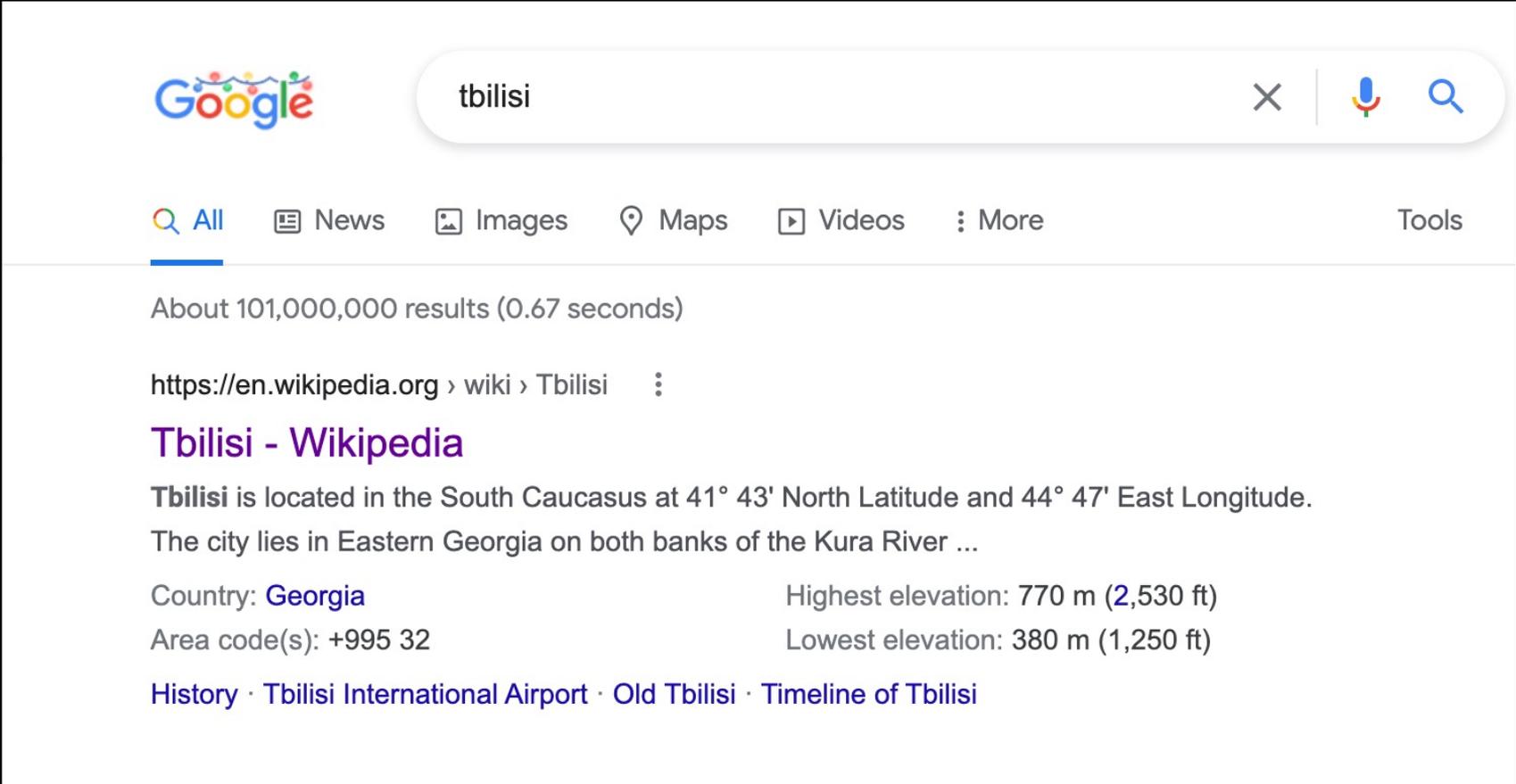
The screenshot displays two product listings side-by-side. The left listing is for a white dress by Norma Kamali, priced at \$158.86. The right listing is for a dark blue dress by Alex Evenings, priced at \$185.00 (originally \$205.00). Both listings include color selection options, star ratings, and Prime eligibility indicators.

Product	Brand	Description	Rating	Price	Prime
	Norma Kamali	Women's Long Sleeve Crew Fishtail Dress to Midcalf	★★★★★ (1)	\$158 ⁸⁶	✓prime
	Alex Evenings	Women's Plus Size Long Dress with Mandarin Neckline Jacket	★★★★☆ (490)	\$185 ⁰⁰ (was \$205.00)	✓prime Prime Try Before You Buy

Ranking features

- Document popularity & reliability
- Where the query word matches
- Query word proximity

tbilisi



The screenshot shows a Google search interface. At the top left is the Google logo. The search bar contains the text 'tbilisi'. Below the search bar are navigation tabs: 'All', 'News', 'Images', 'Maps', 'Videos', 'More', and 'Tools'. The search results show 'About 101,000,000 results (0.67 seconds)'. The first result is from 'https://en.wikipedia.org › wiki › Tbilisi'. The title is 'Tbilisi - Wikipedia'. The snippet reads: 'Tbilisi is located in the South Caucasus at 41° 43' North Latitude and 44° 47' East Longitude. The city lies in Eastern Georgia on both banks of the Kura River ...'. Below the snippet are two columns of information: 'Country: Georgia' and 'Highest elevation: 770 m (2,530 ft)'; 'Area code(s): +995 32' and 'Lowest elevation: 380 m (1,250 ft)'. At the bottom are links for 'History', 'Tbilisi International Airport', 'Old Tbilisi', and 'Timeline of Tbilisi'.

Milk chocolate vs. chocolate milk

The screenshot shows a Google search for "milk chocolate". At the top, the Google logo and search bar are visible with the query "milk chocolate". Below the search bar are navigation tabs for "All", "Shopping", "Images", "News", "Videos", and "More". The search results indicate "About 2,200,000,000 results (0.99 seconds)".

The "Ads · Shop milk chocolate" section features five product listings:

- Milk Chocolates -...**: \$72.75, See's Candi..., Free shipping
- Ghirardelli Milk Chocolate...**: \$209.95, Ghirardelli C..., ★★★★★ (14)
- Milk Chocolate Candies Gift...**: \$25.50, See's Candi...
- Milk Chocolates -...**: \$25.50, See's Candi...
- Godiva Milk Chocolate G...**: \$39.95, Godiva, ★★★★★ (1k+)

Below the ads is a Wikipedia snippet for "Milk chocolate":

https://en.wikipedia.org/wiki/Milk_chocolate

Milk chocolate - Wikipedia

Milk chocolate is manufactured from cocoa, milk and sugar. The ingredient which defines the product as chocolate, cocoa bean, is mainly grown in Southeast Asia, ...

Invented: 1875
Created by: [Daniel Peter](#)
Place of origin: [Switzerland](#)
Main ingredients: [Cocoa butter](#); [Cocoa mas...](#)

[History](#) · [Nutrition](#) · [Production](#) · [Milk chocolate combination bars](#)

On the right side of the search results, there is an image gallery for "Milk chocolate" with a "More images" button. Below the gallery is a detailed information panel for "Milk chocolate" with a share icon:

Milk chocolate

Milk chocolate is a solid chocolate confectionery containing cocoa, sugar and milk. Chocolate was originally sold and consumed as a beverage since pre-Columbian times, and since its introduction to ... [Wikipedia](#)

- Wine pairing
- Ph level
- Gi index
- Temperature

Milk chocolate vs. chocolate milk



chocolate milk



All Images Shopping News Videos More Tools

About 2,230,000,000 results (0.93 seconds)

Ads · Shop chocolate milk



In store

Horizon Organic 1%...

\$6.89

Target

★★★★★ (330)



Slate Milk - Variety 12...

\$35.99

Slate Milk

Free shipping



Horizon Organic...

\$26.67

Amazon.com

Free shipping



In store

Fairlife - Chocolate -...

\$4.99

Safeway



Horizon Organic...

\$7.49

Instacart

★★★★★ (199)

https://en.wikipedia.org/wiki/Chocolate_milk

Chocolate milk - Wikipedia

Chocolate milk is a sweetened chocolate-flavored milk. It can be made by mixing chocolate syrup (or chocolate powder) with milk (from cows, goats, soy, ...

[Separation](#) · [Scientific studies and research](#) · [Nutritional studies](#) · [Nutritional values](#)



More images

Chocolate milk

Drink

Chocolate milk is a sweetened chocolate-flavored milk. It can be made by mixing chocolate syrup with milk. It can be purchased pre-mixed with milk or made at home by blending milk with cocoa powder and a sweetener, melted chocolate, chocolate syrup, or a pre-made powdered chocolate milk mix. [Wikipedia](#)

Types

Serving size

Chemical composition



More images

Milk chocolate

Milk chocolate is a solid chocolate confectionery containing cocoa, sugar and milk. Chocolate was originally sold and consumed as a beverage since pre-Columbian times, and since its introduction to ...

[Wikipedia](#)

Wine pairing

Ph level

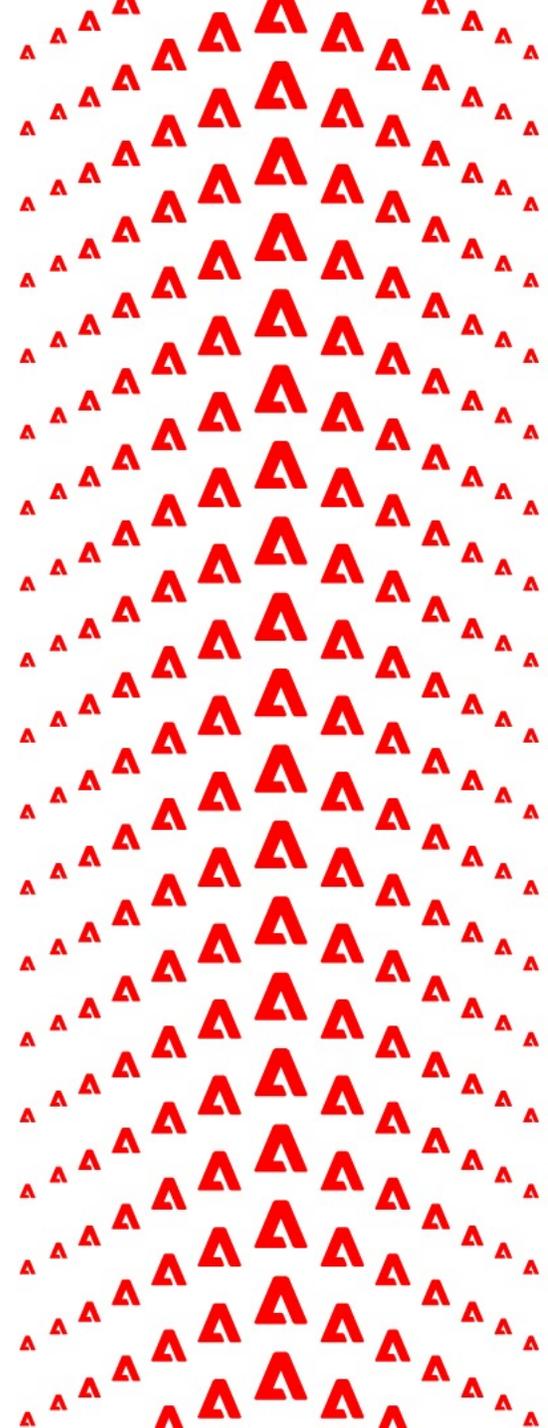
Gi index

Temperature

“How search works” summary

1. For every “document”, find all the words associated with it
2. Build an index with the words and each document that contains them
3. Break the query into words
 1. Lower-casing, lemmatization, stop words, etc.
 2. Meta data
4. Find all the documents associated with each word
5. Return the documents that have all the words of the query
6. Rank them
 1. Behavioral data
 2. Document popularity & reliability
 3. Query word proximity and match location

Why this fails us (or how to break search)



Tricky cases: Vocabulary mismatches

- Synonyms & paraphrases
- Misspellings
- Cross-lingual

- Index vs. query
 - Normalization
 - Expansion

credenza ~ sideboard ~ buffet



SIMPLIHOME Artisan Solid Pine
Wood 54 inch Contemporary
Sideboard Buffet Credenza in Russet
Brown features 2 Doors, 6 Drawers...

Tricky cases: Mapping to structured data

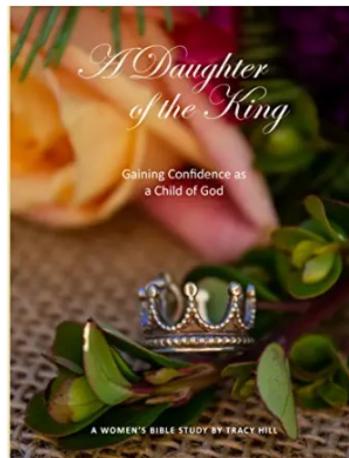
tracy king

- Entities (text):

- brands
- authors

- Special non-text data:

- prices
- dates



The King's Witch: Frances Gorges historical trilogy, Book I
Book 1 of 3: Frances Gorges Historical Trilogy by Tracy Borman | May 14, 2020

★★★★☆ ~ 871

Kindle

\$10⁹⁹ \$15.99

Available instantly

Audible Audiobook

\$0⁰⁰ \$24.95

Free with Audible trial

Available instantly

A Daughter of the King: Gaining Confidence as a Child of God
by Tracy Hill | Jul 7, 2016

★★★★☆ ~ 32

Paperback

\$12⁹⁹

✓prime FREE Delivery Mon, Dec 13

More Buying Choices

\$6.73 (8 used & new offers)

Extra tricky cases

- Negation: “snacks without nuts”
- Syntactic structure:
 - “chocolate milk” vs. “milk chocolate”
 - “who acquired PeopleSoft” vs. “who did PeopleSoft acquire”

 <p>Sponsored ©</p> <p>KIND Bars, Dark Chocolate Nuts & Sea Salt, Gluten Free, 1.4 Ounce Bars, 24 Count</p> <p>24 Count (Pack of 1)</p> <p>★★★★★ ~ 2,632</p> <p>\$28⁴⁴ (\$1.19/Count)</p> <p>\$27.02 with Subscribe & Save discount</p> <p>SNAP EBT eligible</p> <p>✓prime FREE Delivery Sat, Dec 11</p>	 <p>Macadamia Nut Butter 100% Pure Raw and Without Any Additives or Sugar Made from 100% Macadamia Nuts Perfect Keto Diet Snack - 12 Oz</p> <p>12 Ounce (Pack of 1)</p> <p>★★★★★ ~ 130</p> <p>\$29⁴⁹ (\$2.46/Ounce)</p> <p>✓prime FREE Delivery Sun, Dec 12</p> <p>Only 6 left in stock - order soon.</p> <p>More Buying Choices</p> <p>\$24.45 (2 new offers)</p>	 <p>Without A Trace Top 9 Allergen-Friendly, Nutrient Dense, Plant Based Power Bites - Nut Free Snacks, Gluten Free And Snack Size. Convenient &...</p> <p>★★★★★ ~ 52</p> <p>\$26⁹⁰ (\$2.13/Ounce)</p> <p>✓prime FREE Delivery Sun, Dec 12</p>	 <p>Deluxe Roasted Salted Mixed Nuts in Sea Salt by Its Delish – Bulk 2 lbs - Premium Roasted Nut Mix Without Peanuts – Delicious Snacking Nuts –...</p> <p>★★★★★ ~ 52</p> <p>\$24⁹⁹ (\$24.99/Count)</p> <p>Get it Fri, Dec 17 - Tue, Dec 21</p> <p>FREE Shipping</p>	 <p>Peanut Free & Tree Nut Free Snack Sack - Back to School Snack Variety Pack or College Care Package Popcorn, Chips, Cookies, & Sweets –...</p> <p>42 Piece Set</p> <p>★★★★★ ~ 20</p> <p>Save 5%</p> <p>\$34⁹⁵ (\$34.95/Count) \$36.95</p> <p>Save more with Subscribe & Save</p> <p>Lowest price in 30 days</p> <p>✓prime FREE Delivery Sun, Dec 12</p>	 <p>Made Good Healthy Snacks Variety Pack, 40 Ct - Organic Assortment of Granola Bars, Granola Mini Snack Packs, Crispy Squares; Individually...</p> <p>7 Piece Assortment</p> <p>★★★★★ ~ 2,502</p> <p>\$29⁹⁹ (\$29.99/Count)</p> <p>\$28.49 with Subscribe & Save discount</p> <p>✓prime FREE Delivery Sat, Dec 11</p> <p>Climate Pledge Friendly</p> <p>See 1 certification</p>	 <p>Allergy Friendly GLUTEN FREE and VEGAN (15Ct): DAIRY, WHEAT, PEANUT, TREE NUT, SOY, and EGG FREE) Healthy Snacks Care Package...</p> <p>15 Piece Set</p> <p>★★★★★ ~ 6</p> <p>\$39⁹⁵ (\$2.66/Count)</p> <p>Save more with Subscribe & Save</p> <p>✓prime FREE Delivery Sat, Dec 11</p>	 <p>Enjoy Life Foods Seed and Fruit Mix, Peanut Free Trail Mix, Mountain Mambo, 39.12 Ounce (Pack of 24)</p> <p>1.63 Ounce (Pack of 24)</p> <p>★★★★★ ~ 1,099</p> <p>\$32⁹⁸ (\$0.84/Ounce)</p> <p>\$31.33 with Subscribe & Save discount</p> <p>40% off your first subscription order</p> <p>SNAP EBT eligible</p> <p>✓prime FREE Delivery Sat, Dec 11</p> <p>More Buying Choices</p> <p>\$32.50 (6 new offers)</p>
---	--	---	--	---	--	--	---

Robustness: Embeddings & deep learning

- Encode everything in a high dimensional space
- Similar words (or images) close to one another
- Different modalities (text & images) can map to the same space

Robustness: Embeddings & deep learning



St Adobe Stock Images ▾

Find Similar Clear



All attributes

Content

Click a dot to search by object

Search by position & size

Color

Composition

Embeddings: Where they work well

- Vocab mismatches between queries and items
- Misspellings
- Mapping similar queries together
- Simple multi-modal (text-to-image) search
- Stuff everyone is looking for
 - Especially when combined with behavioral & other ranking signals

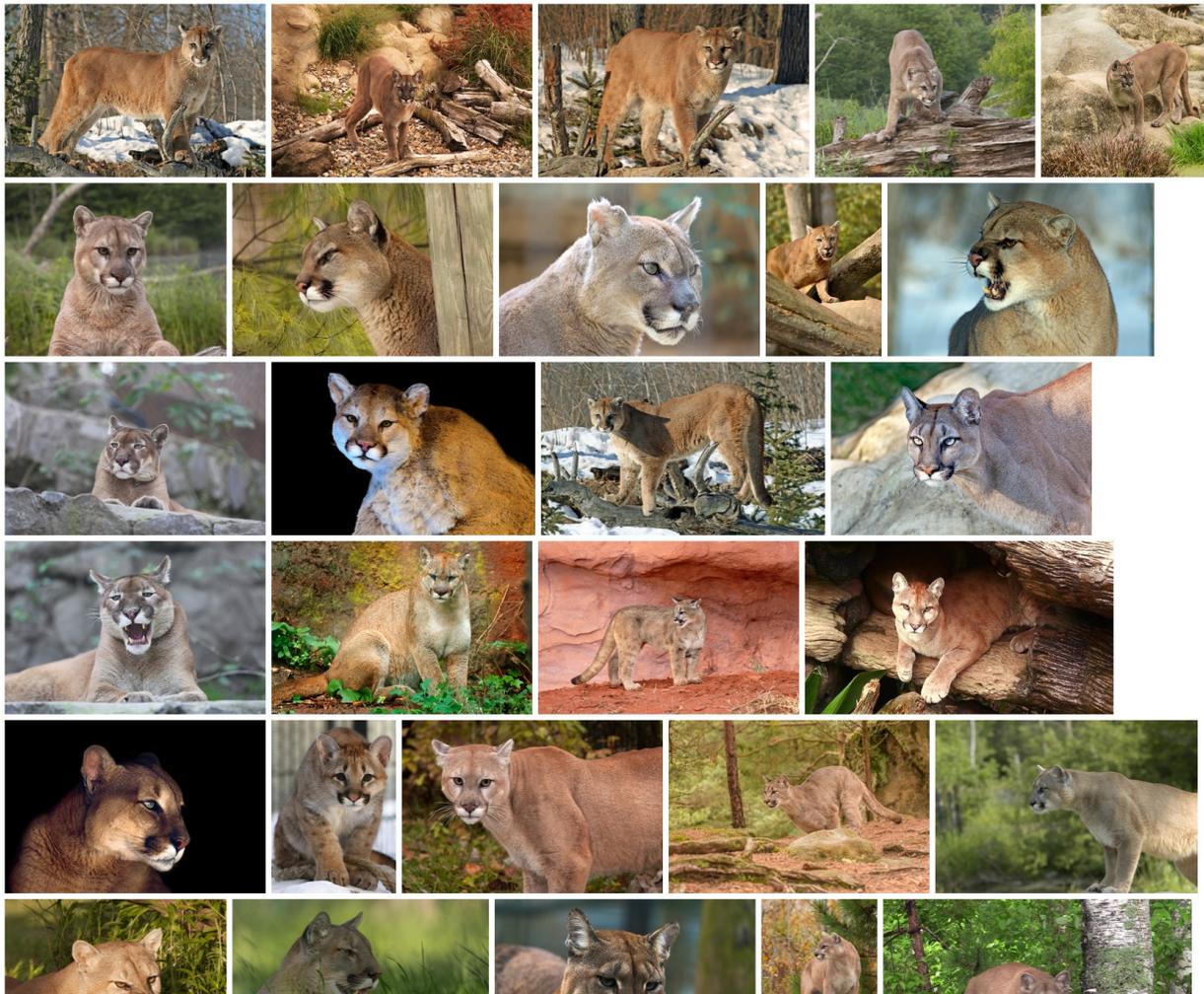
credenza ~ sideboard ~ buffet



SIMPLIHOME Artisan Solid Pine Wood 54 inch Contemporary Sideboard Buffet Credenza in Russet Brown features 2 Doors, 6 Drawers...

cougar ~ puma ~ mountain lion

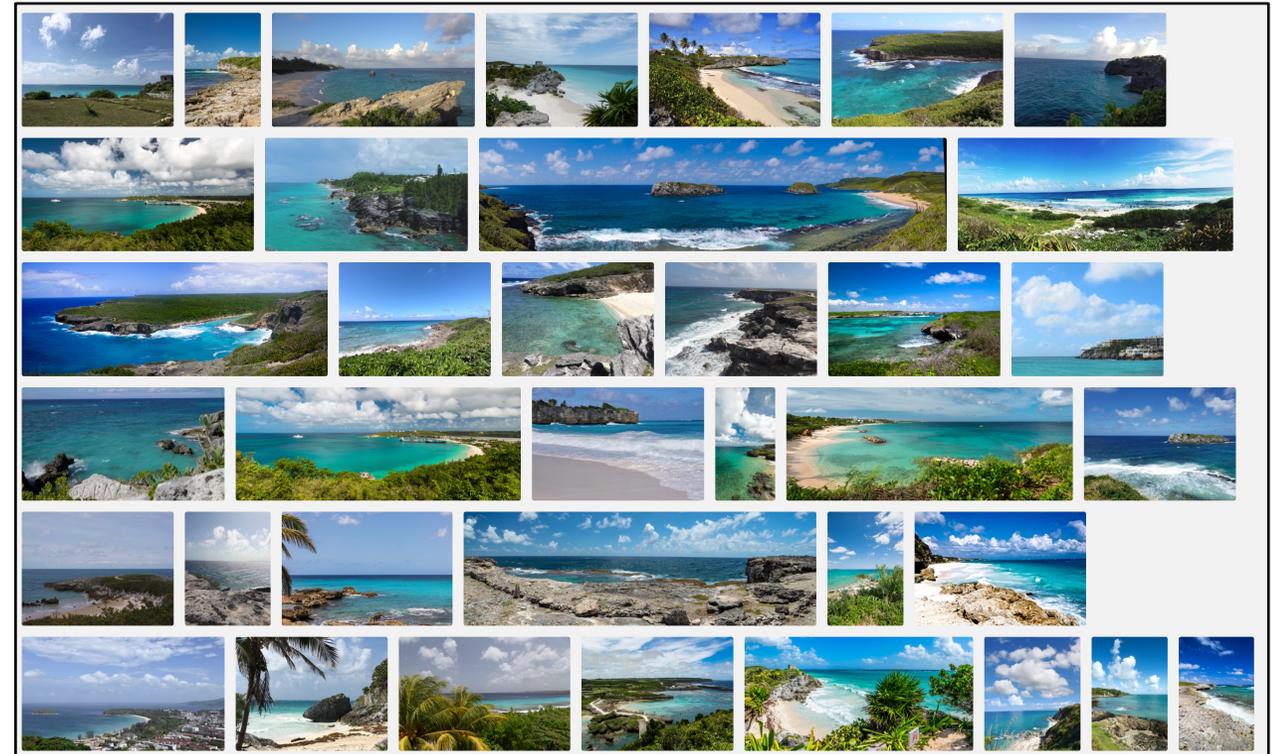
cougar



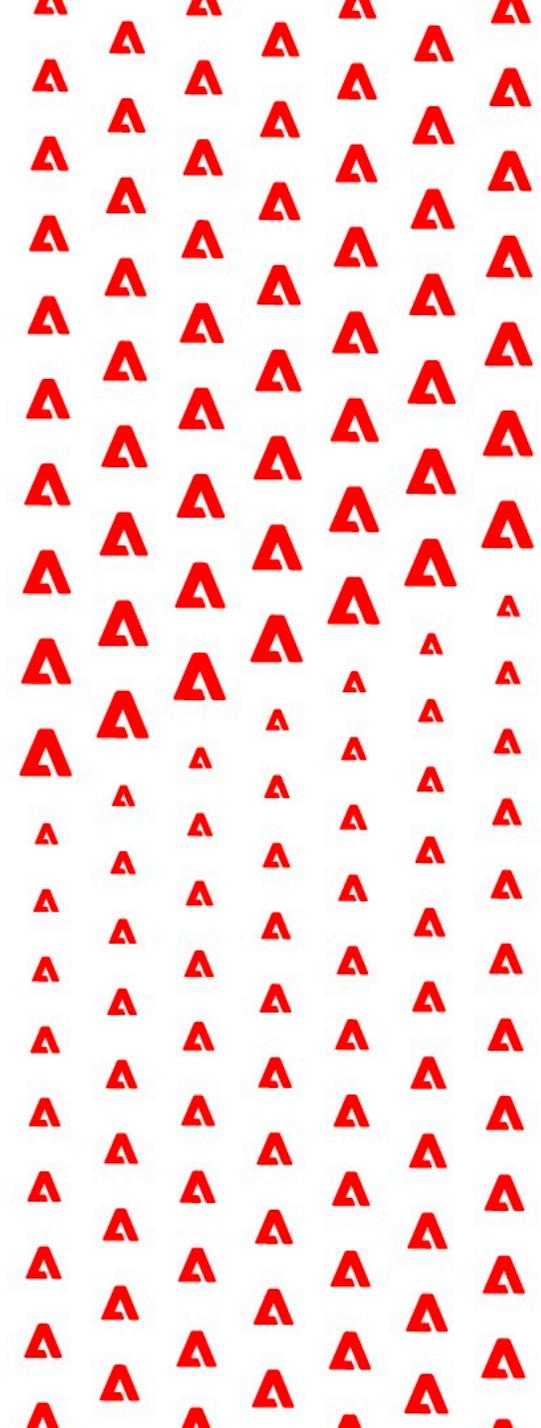
Embeddings: Where they fail

- Things that distribute similarly but with different meanings
 - numbers, model numbers
 - people names of same gender & ethnicity
- Negation & counterfactuals
- Information determined by syntactic structure (not just co-occurrence)

Bermuda coast



Towards a solution: Adding (and removing) structure



Structure where it matters

- Why not everywhere:
 - costly: index size, query processing
 - specialized knowledge to create
- General: Entity recognition

The image shows a Google search for 'montpellier'. The search bar at the top contains the word 'montpellier' and is highlighted with a blue box. Below the search bar, there are navigation links for 'All', 'Images', 'Maps', 'News', 'Videos', and 'Tools'. The search results show 'About 2,630,000,000 results (0.98 seconds)'. The first result is 'Montpellier - Wikipedia', with a snippet: 'Montpellier is the third-largest French city near the Mediterranean coast, behind Marseille and Nice. It is the seventh-largest city of France, ...'. Below this, there are details for the region (Occitanie), area (56.88 km2), population (290,053), and department (Hérault). The second result is 'Montpellier HSC - Wikipedia', with a snippet: 'Montpellier Hérault Sport Club commonly referred to as Montpellier HSC or simply Montpellier, is a French professional football club based in the city of ...'. Below this, there are details for the league (Ligue 1), nickname (La Paillade), and head coach (Olivier Dall'Oglio). There is also a 'People also ask' section with four questions: 'What is Montpellier famous for?', 'Is Montpellier worth visiting?', 'Is Montpellier France expensive?', and 'Is Montpellier good place to live?'. At the bottom, there is a 'Top stories' section. On the right side, there is a knowledge panel for 'Montpellier' with a photo of the city and a map. The panel includes the text 'City in France' and a detailed description: 'Montpellier is a city in southern France, 10km inland from the coast of the Mediterranean Sea. The town's stately Gothic Cathédrale Saint-Pierre, distinguished by conical towers, dates to 1364. The city's Antigone district is a chic, modern development inspired by neoclassical motifs. Paintings from French and European Old Masters hang at the Musée Fabre. — Google'. Below the description, there are statistics: 'Area: 21.96 mi²', 'Elevation: 89'', 'Weather: 47°F (8°C), Wind W at 4 mph (6 km/h), 78% Humidity weather.com', 'Local time: Monday 6:53 PM', and 'Population: 277,639 (2015) INSEE'. At the bottom of the panel, there is a 'Plan a trip' section with a camera icon and the text 'Things to do'.

Structure where it matters

- Why not everywhere:
 - costly: index size, query processing
 - specialized knowledge to create
- General: Entity recognition
- Ecommerce: Prices + over/under/from/to

Google montpellier

montpellier

dresses over \$100

Norma Kamali
Women's Long Sleeve Crew Fishtail Dress to Midcalf
★★★★★ 1
\$158⁸⁶
prime

Alex Evenings
Women's Plus Size Long Dress with Mandarin Neckline Jacket
★★★★☆ 490
\$185⁰⁰ \$205.00
prime
Prime Try Before You Buy

ance, 10km inland
n Sea. The town's
erre, distinguished
The city's Antigone
ent inspired by
French and
Musée Fabre.

mph (6 km/h), 78%

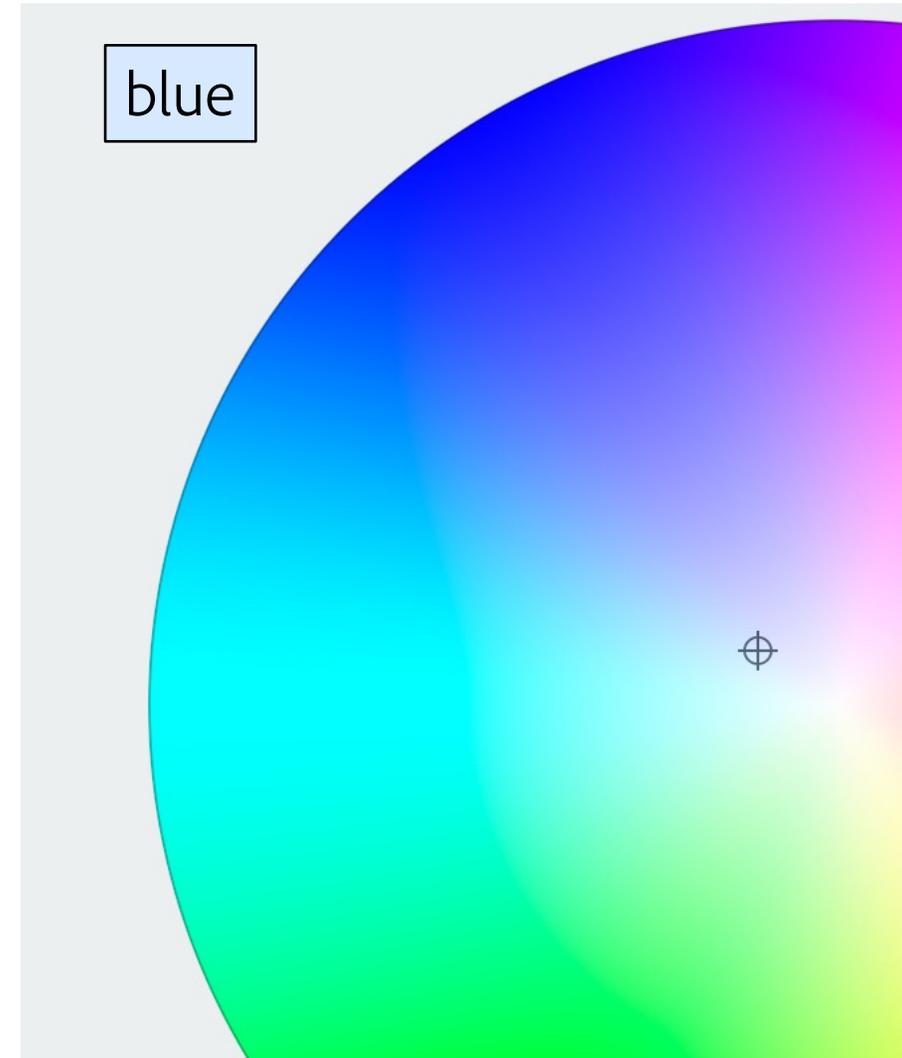
Structure where it matters

- Why not everywhere:
 - costly: index size, query processing
 - specialized knowledge to create
- General: Entity recognition
- Ecommerce: Prices + over/under/from/to
- Images: Color-object queries



Robustness where you need it

- Allow for looser matching when needed
- Synonyms and misspellings
- Null & low result recovery
- Color terms

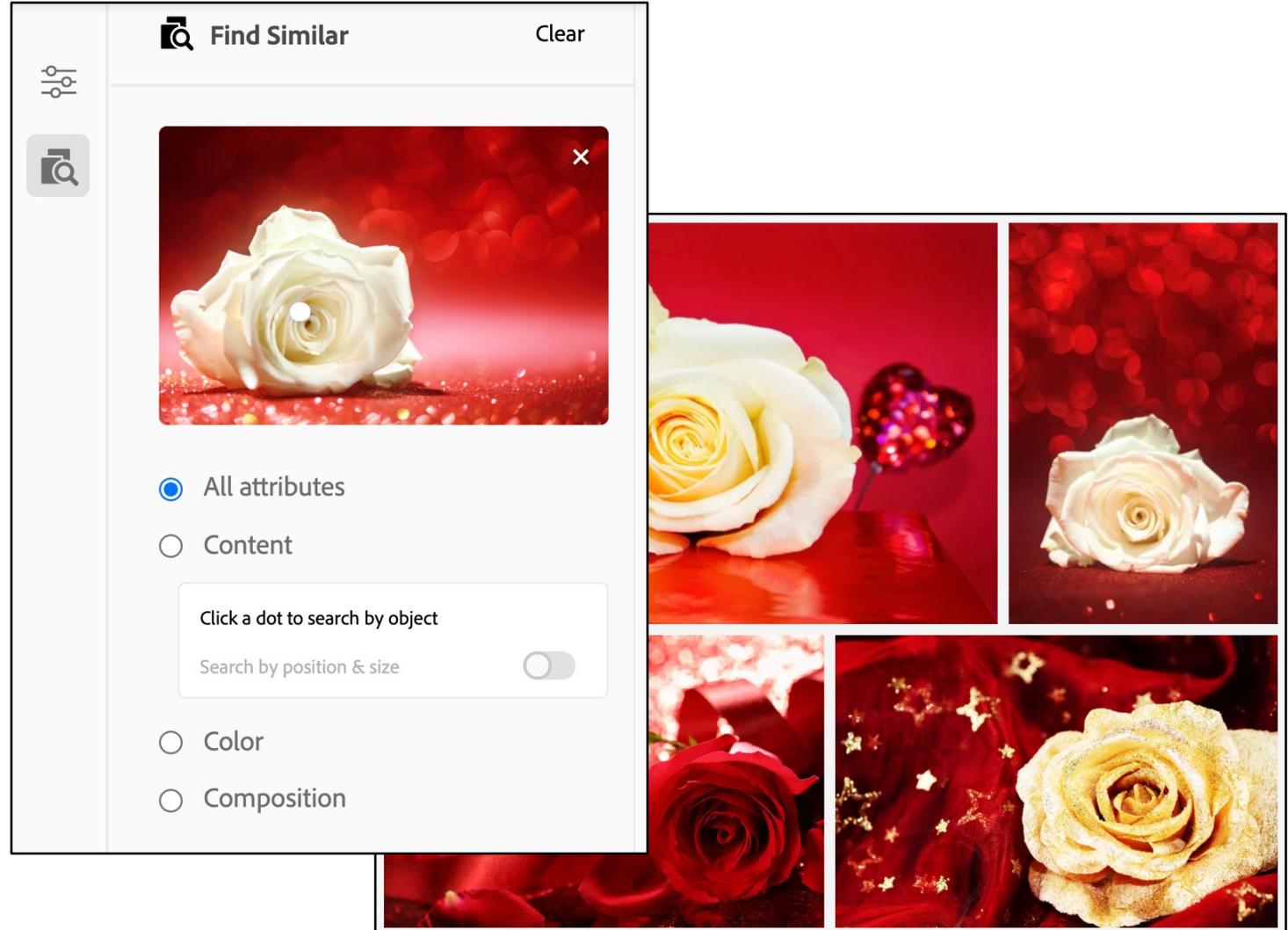


Adobe Stock color-object queries: white rose red background



Adobe Stock color-object queries

- The issue:
 - Long tail of color-object queries
 - Visually jarring
 - No easy way to recover
- The cause:
 - Images with “bag of tags” & titles
 - Red roses & white backgrounds are much more popular



Solution: Specialized Structured Data

- Images:
 - Identify prominent objects (+ background)
 - Identify their colors
 - Store in the index
- Query:
 - Identify color terms
 - Identify the objects they modify
- Matching and ranking:
 - Aggressive solution: Only return images with the same color and object as the query
 - Conservative solution: Rank higher images with the same color and object as the query

Color-object queries

- Named entity recognition model for colors
 - Which colors to focus on: kindergarten colors; modified (“hot pink toaster”); rarer (“salmon background”)
 - False positives: ethnicities “white nurse”; general “snow white”; style “black & white”
 - Complex cases: coordination (“red and green peppers”)
- Dependency parser to detect object
- Custom-trained, short-text versions
- Language specific



Color-object images

- Find and mask the objects
 - Label them
 - List of objects from queries
 - Auto-tagger confidence
- Determine their color
- Background as a special case



Object: cup & Color: white
Object: cookie & Color: brown
Object: background & Color: blue

Color-object issues

- What are the colors?
 - Image color embeddings (language independent)
 - Query text2color model (language dependent)
- What are the objects?
 - Index language-independent concepts
 - Query text-to-concept mappings (language dependent)



Object: cnx:123:cup & color: []

Object: cnx:456:cookies & color: []

Object: cnx:0:background & color: []

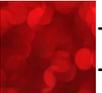
Color-object in action

- Query color NER + dependency parse + text2color + object text to concept
- Query: object concept + color embedding (+ any other keywords)
- Matching & ranking
 - Images with the object concept (+ any other keywords)
 - Image-object & query color embedding similarity calculation



white rose red background

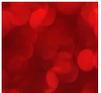
object=cnx:789:rose & color = []

object=cnx:0:background & color = []

images with cnx:789:rose

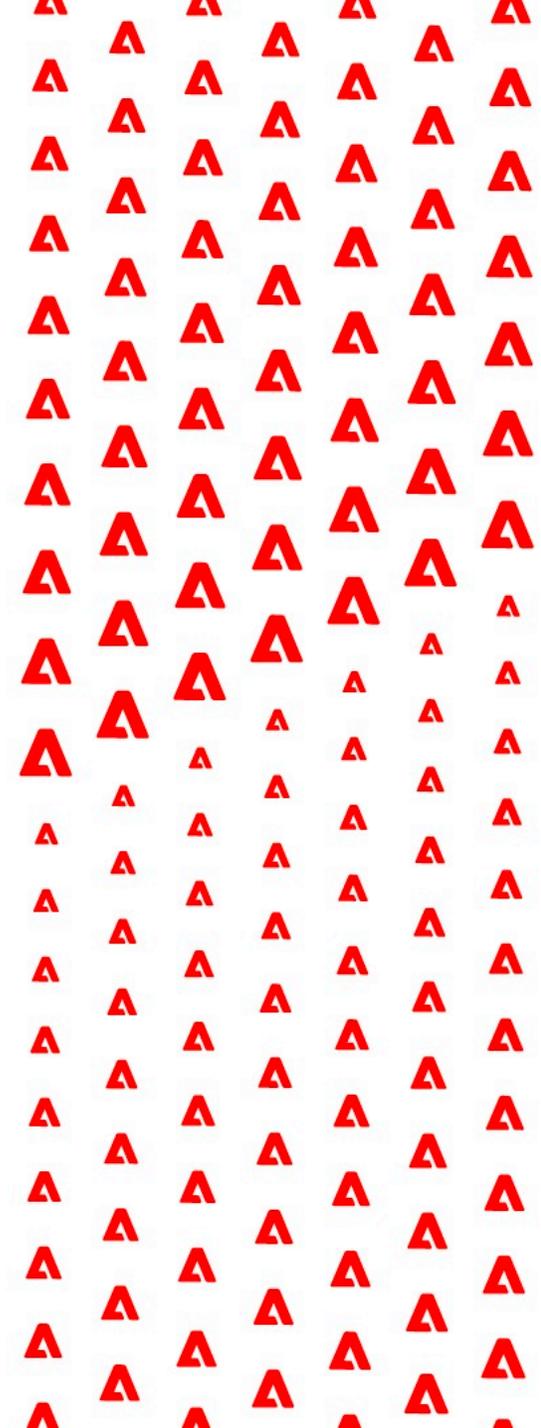


images with

cnx:0:background & color = []

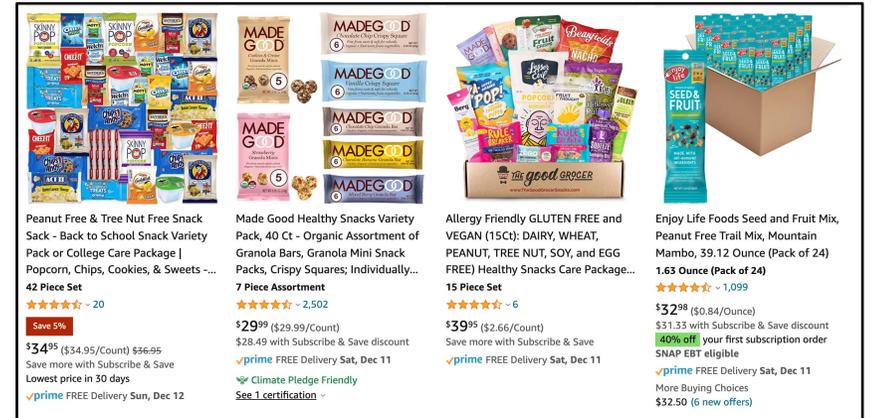
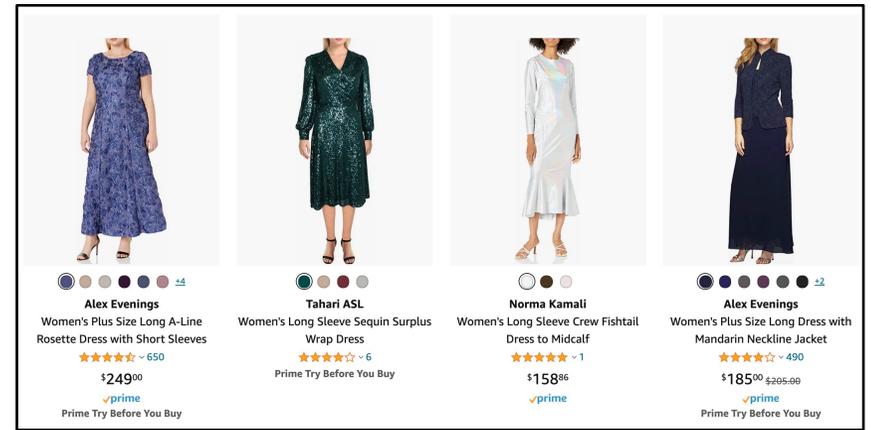
cnx:789:rose & color = []

Wrapping up: Where are we?



Structuring Search

- Search historically depends on inverted indices and behavioral ranking signals
- Embeddings and deep learning help with robustness
- Structured representations are needed for deeper understanding
 - High-value queries
 - Requires knowledge to build structure
 - Indexing structure can be more costly
- Future of search is hybrid: Structure + robustness



Questions?

