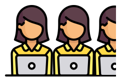
Towards Low-Code/No-Code Insight Discovery

The First EPIC Advance
October 26, 2022

Aditya Parameswaran

A Matter of Scale

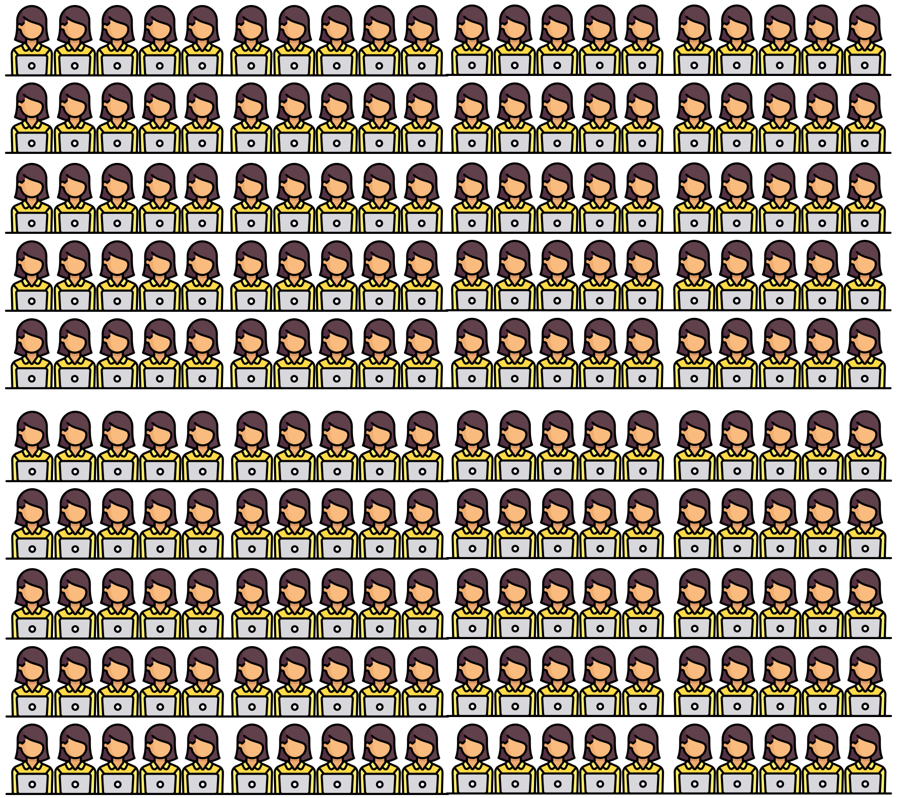
Programmers
20 M



Data Analysts
50 M

Spreadsheet Users
1 B = 1000M

Target of
Low Code Tools



Search Users
4 B = 4000M

Target of
No Code Tools

How do we make

training an ML model

building a dashboard

identifying patterns

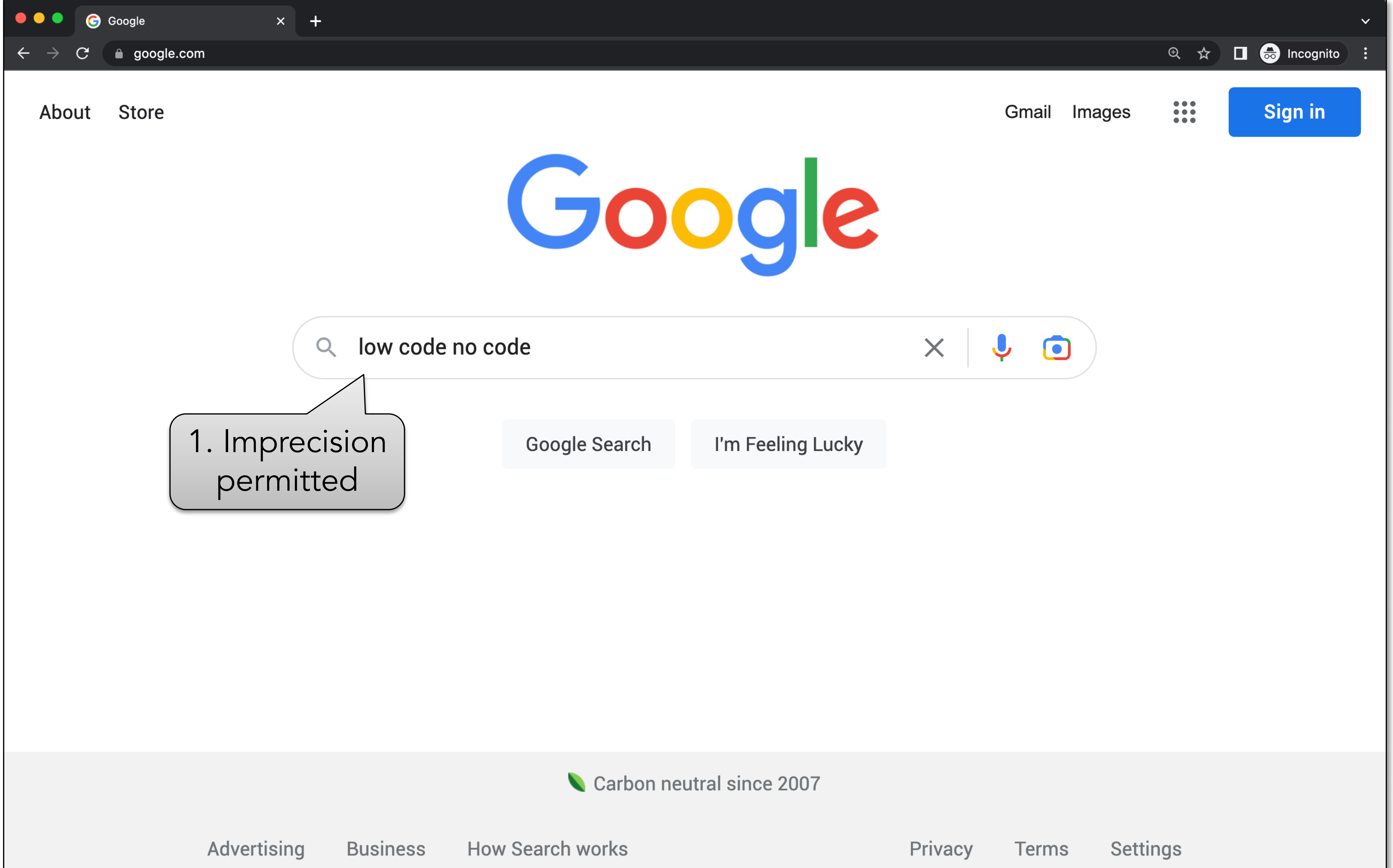
data cleaning

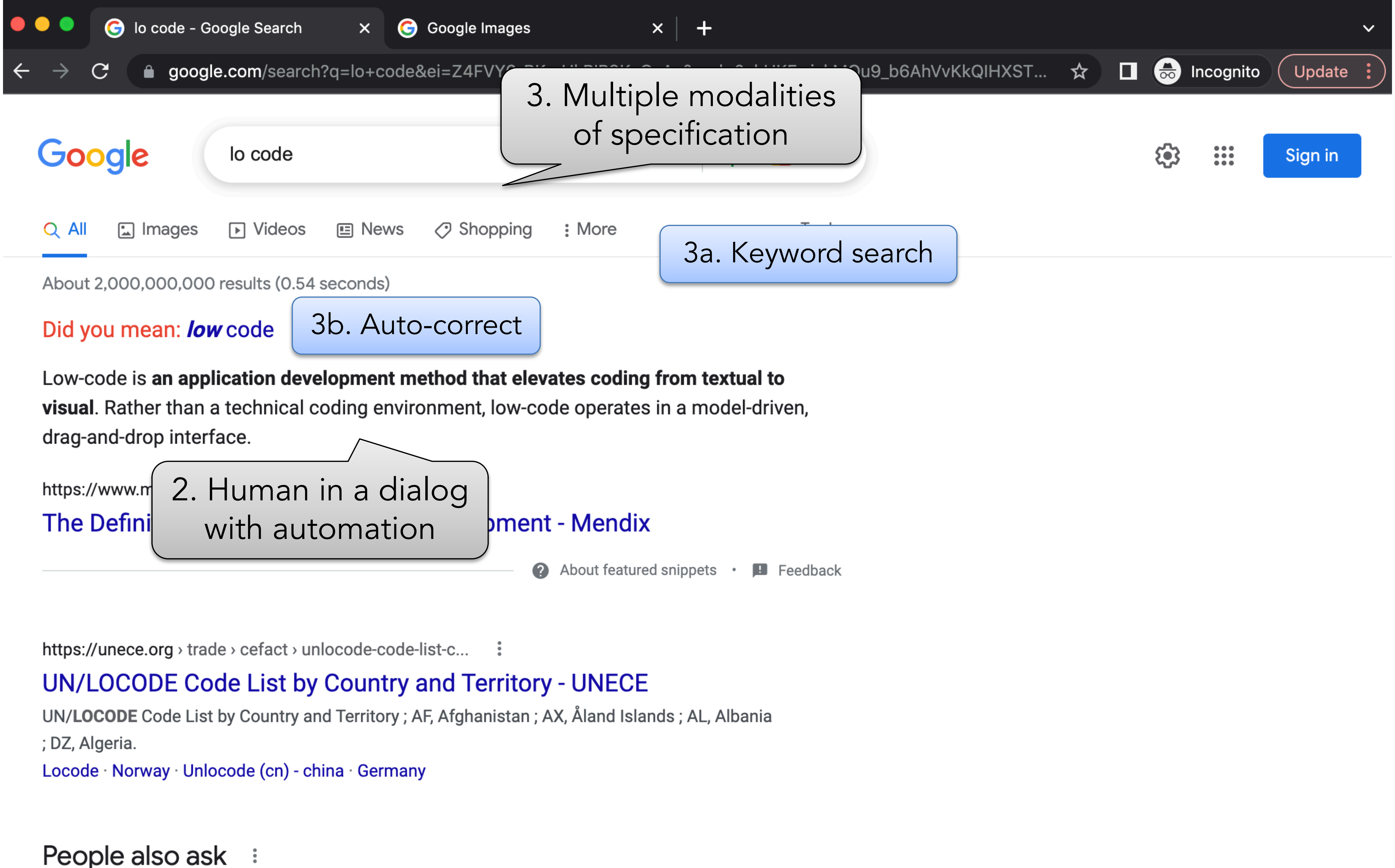
…

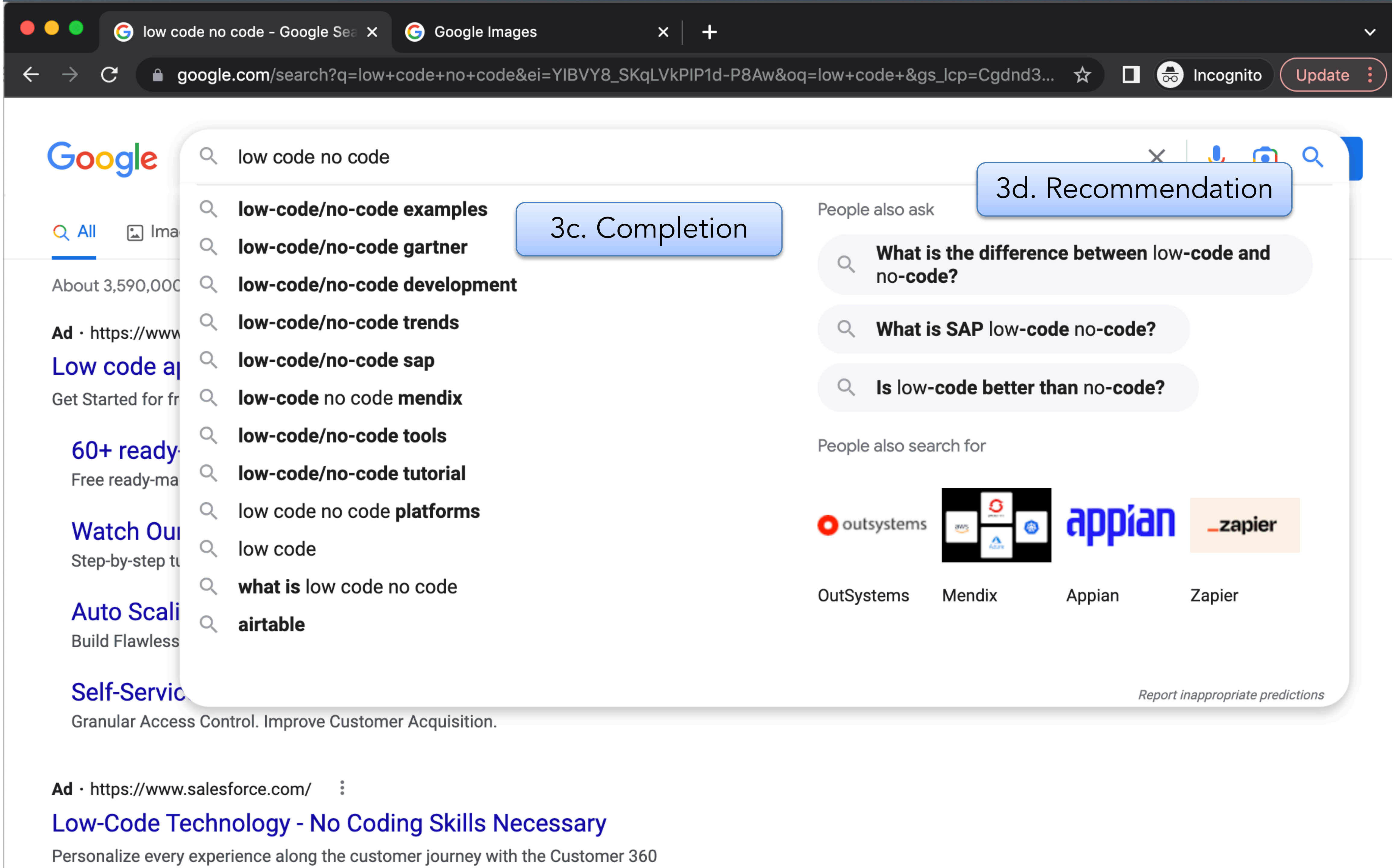
$\underbrace{\qquad\qquad}_{\text{Insight Discovery}}$

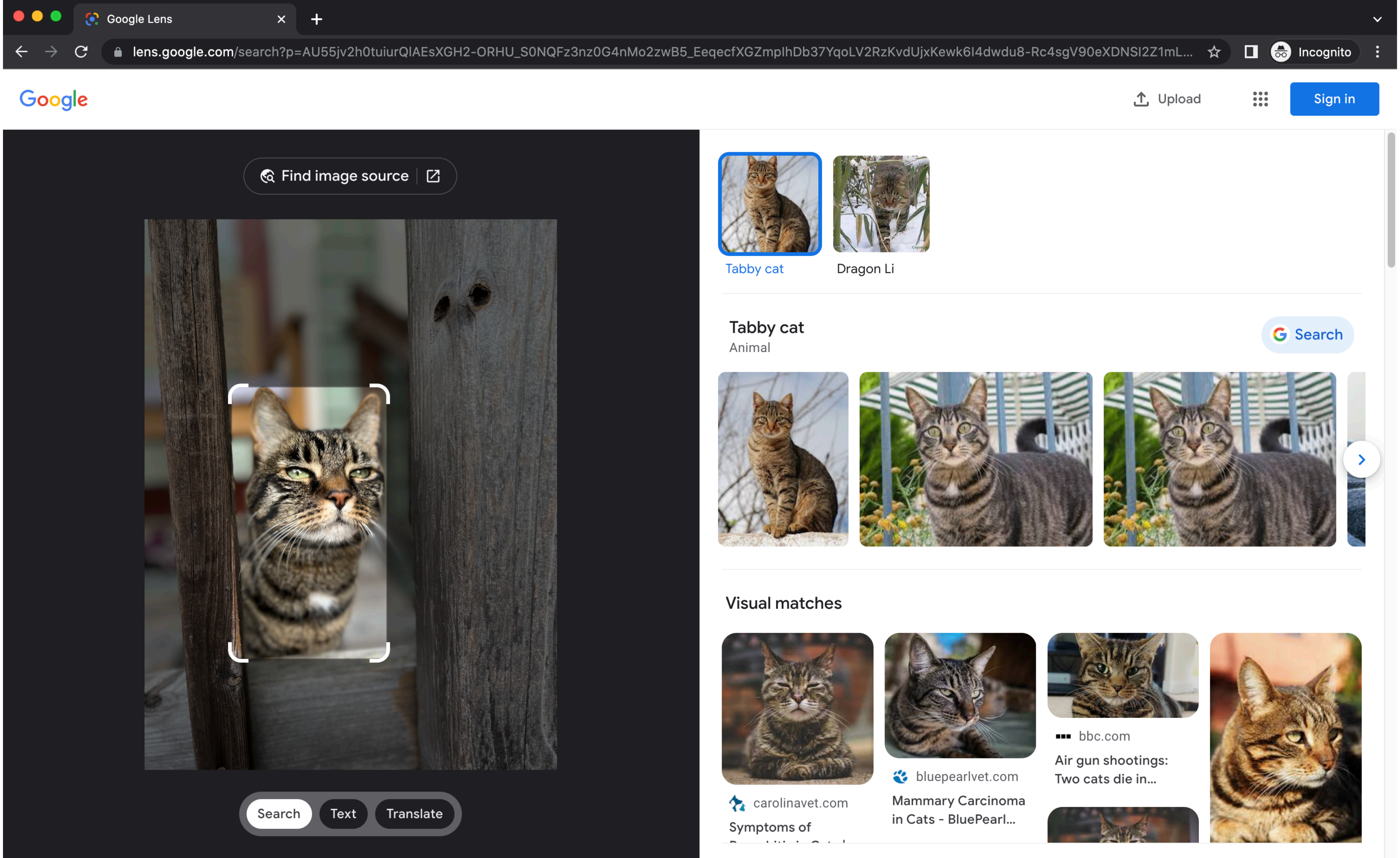
as easy as **search**?

So, what's so great about search?


Google
google.com
Incognito
About
Store
Gmail
Images
Sign in
Google
low code no code
1. Imprecision permitted
Google Search
I'm Feeling Lucky
Carbon neutral since 2007
Advertising
Business
How Search works
Privacy
Terms
Settings


lo code - Google Search
Google Images
google.com/search?q=lo+code&ei=Z4FVY
u9_b6AhVvKkQIHXST...
Incognito
Update
3. Multiple modalities of specification
Google
lo code
Sign in
All
Images
Videos
News
Shopping
More
3a. Keyword search
About 2,000,000,000 results (0.54 seconds)
Did you mean: low code
3b. Auto-correct
Low-code is an application development method that elevates coding from textual to visual. Rather than a technical coding environment, low-code operates in a model-driven, drag-and-drop interface.
https://www.m
2. Human in a dialog with automation
The Defini
pment - Mendix
About featured snippets
Feedback
https://unece.org › trade › cefact › unlocode-code-list-c...
UN/LOCODE Code List by Country and Territory - UNECE
UN/LOCODE Code List by Country and Territory ; AF, Afghanistan ; AX, Åland Islands ; AL, Albania ; DZ, Algeria.
Locode · Norway · Unlocode (cn) - china · Germany
People also ask


low code no code - Google Search
Google Images
google.com/search?q=low+code+no+code&ei=YIBVY8_SKqLVkPIP1d-P8Aw&oq=low+code+&gs_lcp=Cgdnd3...
Incognito
Update
Google
low code no code
3d. Recommendation
3c. Completion
low-code/no-code examples
low-code/no-code gartner
low-code/no-code development
low-code/no-code trends
low-code/no-code sap
low-code no code mendix
low-code/no-code tools
low-code/no-code tutorial
low code no code platforms
low code
what is low code no code
airtable
People also ask
What is the difference between low-code and no-code?
What is SAP low-code no-code?
Is low-code better than no-code?
People also search for
OutSystems
Mendix
Appian
Zapier
Report inappropriate predictions
All
Ima
About 3,590,000
Ad · https://www
Low code a
Get Started for fr
60+ ready
Free ready-ma
Watch Ou
Step-by-step t
Auto Scali
Build Flawless
Self-Servic
Granular Access Control. Improve Customer Acquisition.
Ad · https://www.salesforce.com/
Low-Code Technology - No Coding Skills Necessary
Personalize every experience along the customer journey with the Customer 360


Google Lens
lens.google.com/search?p=AU55jv2h0tuiurQIAEsXGH2-ORHU_S0NQFz3nz0G4nMo2zwB5_EeqecfXGZmplhDb37YqoLV2RzKvdUjxKewk6I4dwdu8-Rc4sgV90eXDNSI2Z1mL...
Incognito
Google
Upload
Sign in
Find image source
Search
Text
Translate
Tabby cat
Dragon Li
Tabby cat
Animal
Search
Visual matches
carolinavet.com
Symptoms of
bluepearlvet.com
Mammary Carcinoma in Cats - BluePearl...
bbc.com
Air gun shootings: Two cats die in...

Key Design Principles from Search

1. Imprecision tolerated
   – fuzziness, incompleteness, inaccuracies, …

2. Human in a dialog with automation
   – both doing what they do best

3. Multiple specification modalities
   – examples, keyword search, interactions/selection

Adapting the Design Principles for Data Work

How do we make *training an ML model* / *building a dashboard* / *identifying patterns* / *data cleaning* / … as easy as **search**?

- What input specifications does the system support?
- How does the system engage in a dialog with the user?
- How does the system deal with imprecision (aka many matches)?
- How does the system return results in an interactive manner?

Case Study: Visual Query Systems

From Sketching to Natural Language: Expressive Visual Querying for Accelerating Insight. CACM, 2022

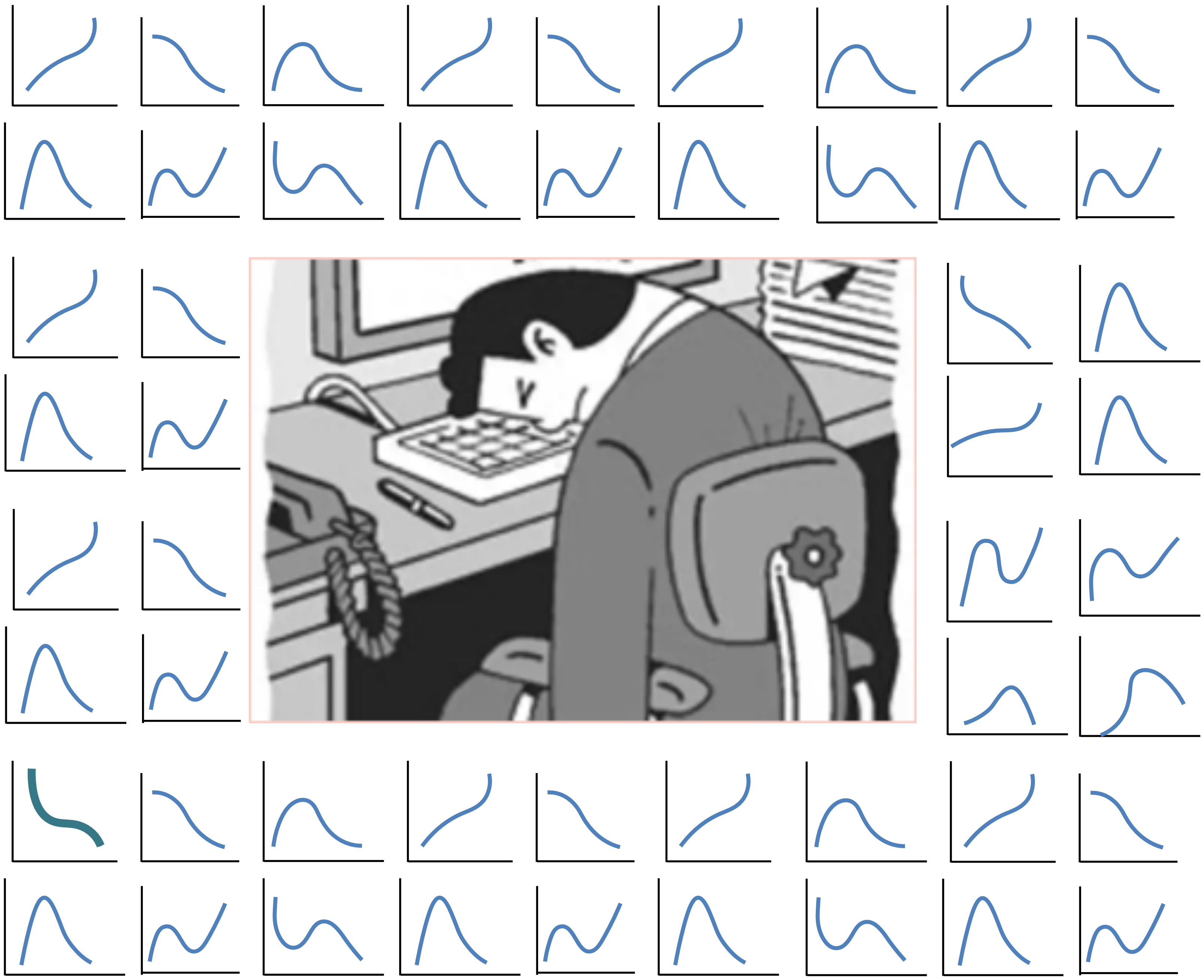
ShapeSearch: A Flexible and Efficient System for Shape-based Exploration of Trendlines. SIGMOD 2020 [Best Paper Award]

You can't always sketch what you want: Understanding Sensemaking in Visual Query Systems. VIS 2019

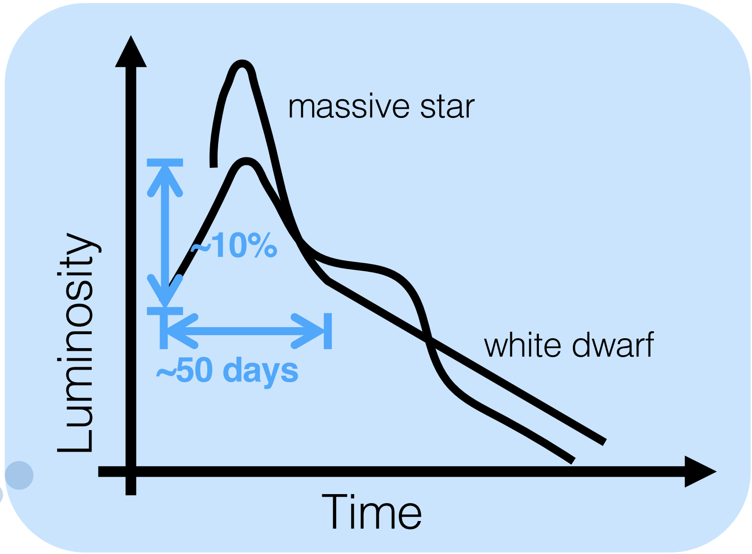
Effortless Visual Data Exploration with Zenvisage: An Interactive and Expressive Visual Analytics System. VLDB 2017

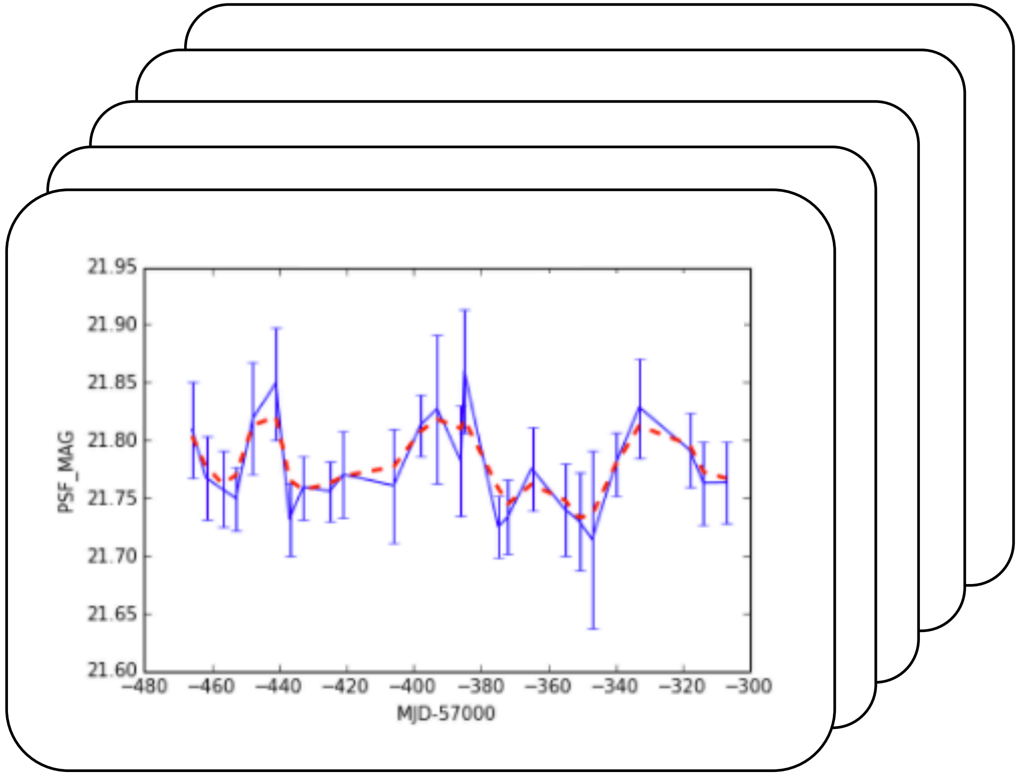
Analysts often have a need to search for patterns in a set of line charts

Often too many charts to look at!



Concrete Example: Astronomy





1000s of light curves to examine!

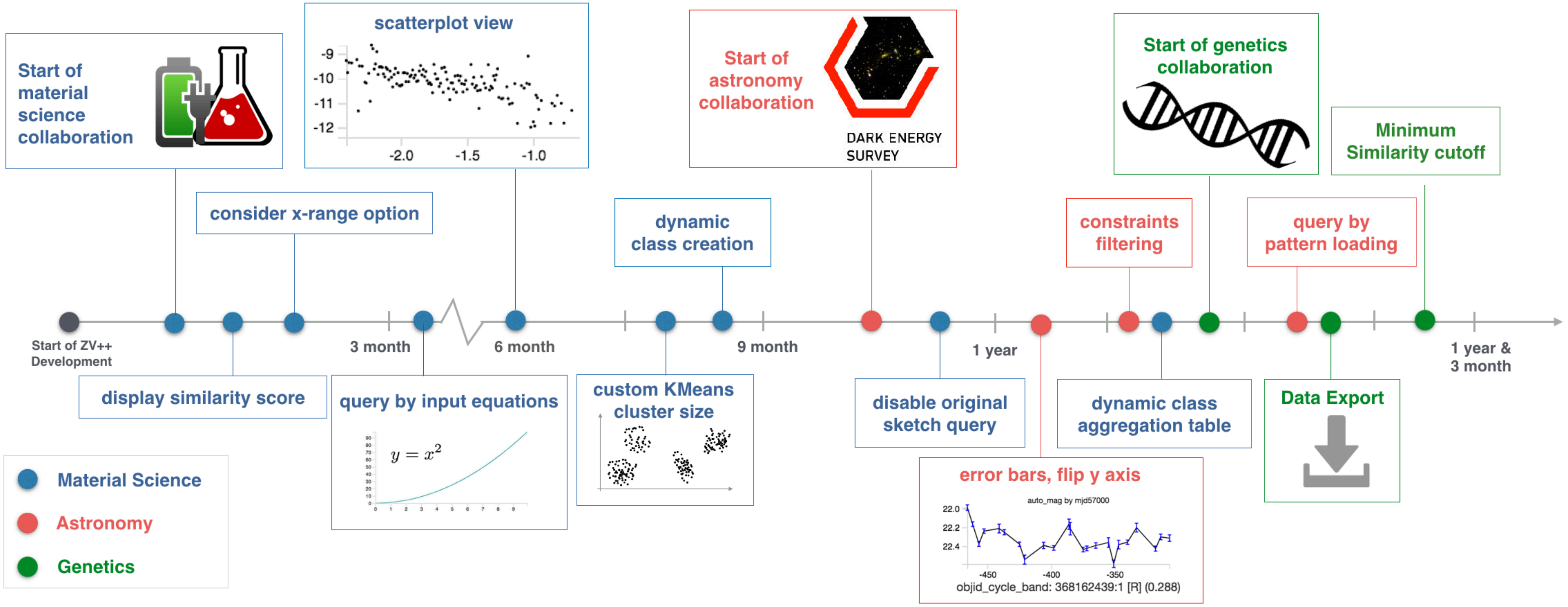
Why not use Visual Query Systems?

TimeSearcher [Hochheiser & Shneiderman 2004]
QuerySketch [Wattenberg 2001]
QueryLines [Ryall et al. 2005]
SoftSelect [Holz & Feiner 2009]
Google Correlate [Mohebbi 2011]

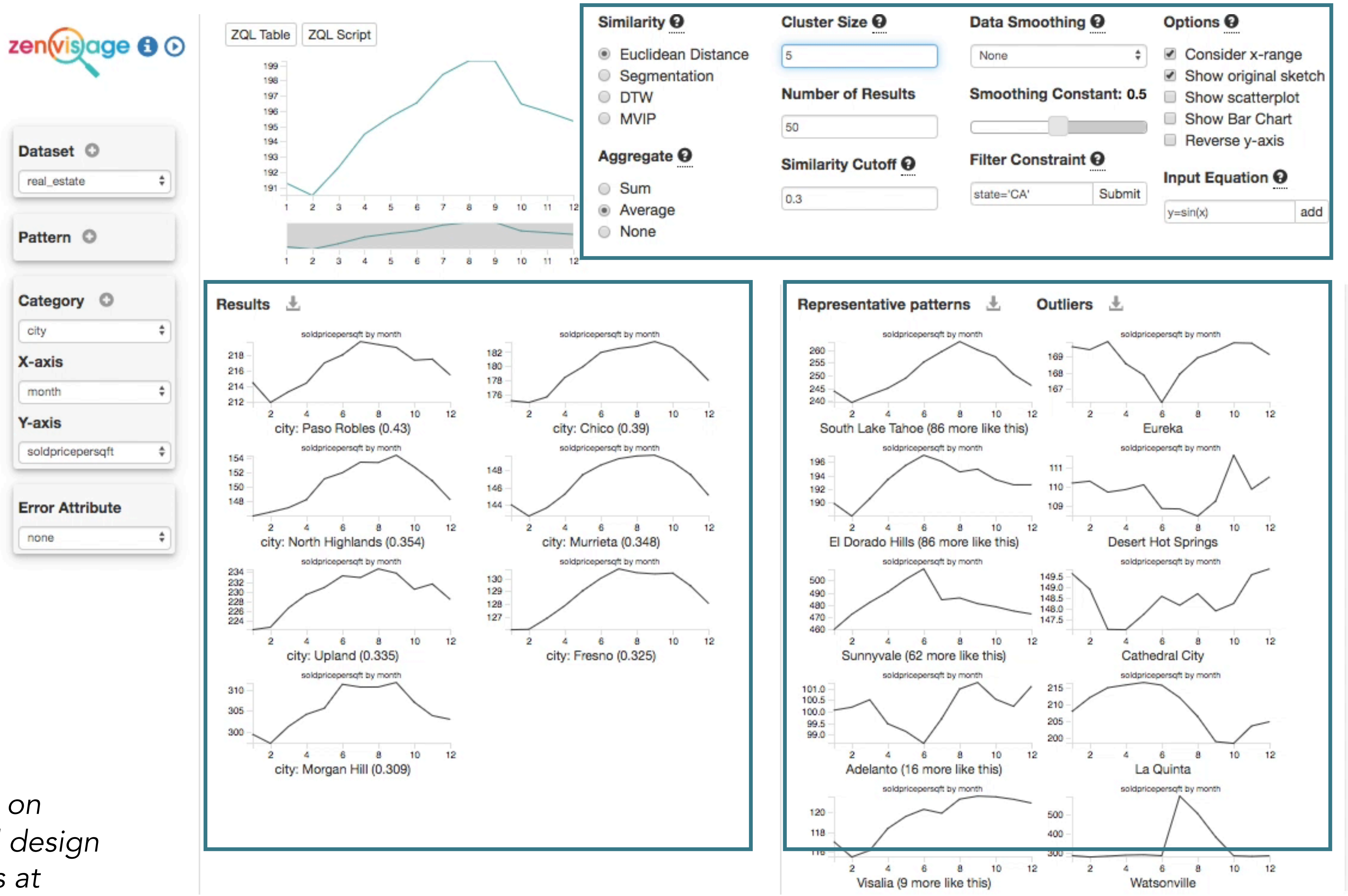
TimeSketch [Eichmann & Zgraggen 2015]
SketchQuery [Correll & Gleicher 2016]
Zenvisage [Siddiqui et al 2017]
Qetch [Mannino & Abouzied 2018]

However, none of this prior work has seen adoption!

Can we develop a “useful” visual query system?

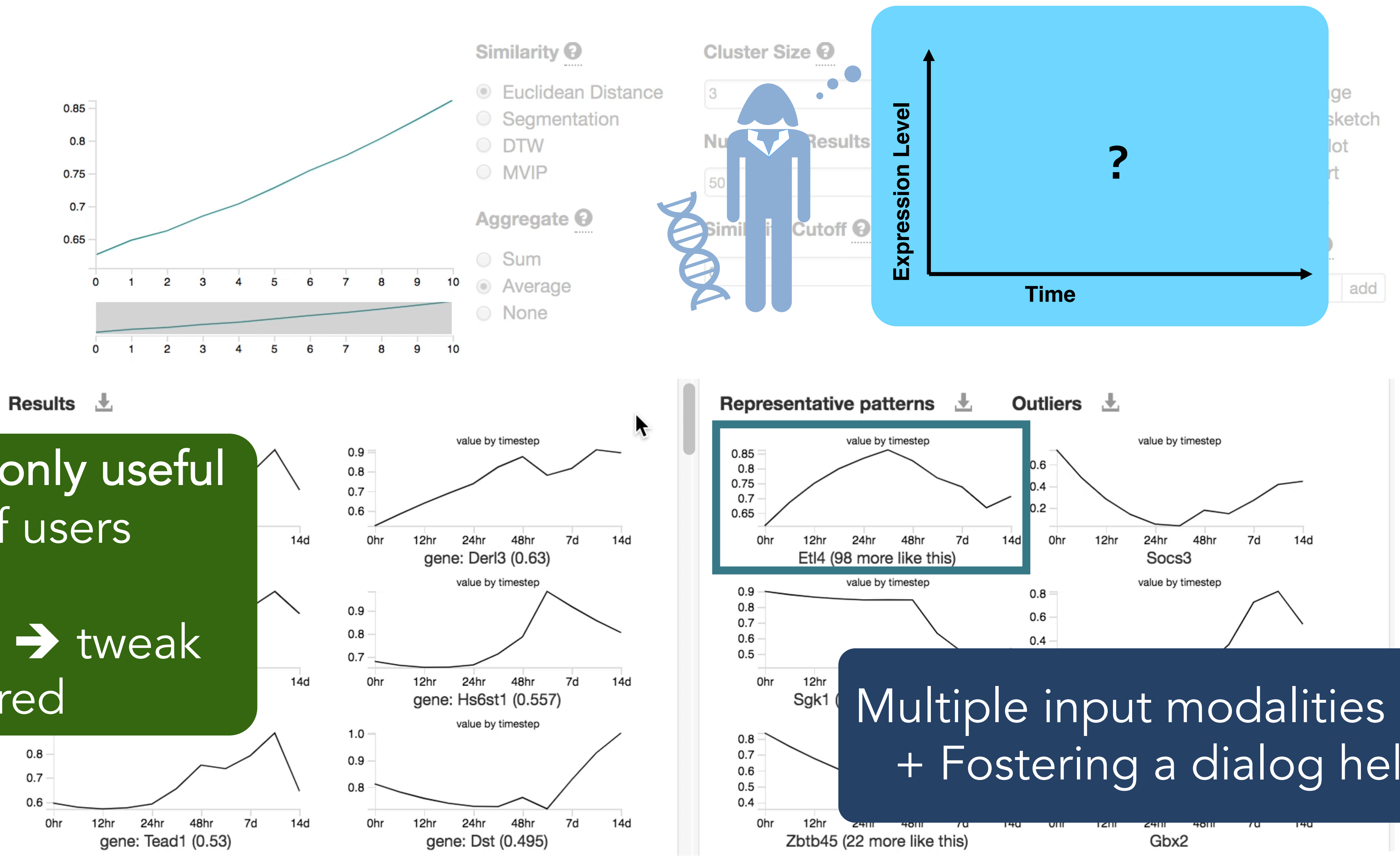


Year-long design with battery scientists, astronomers, and genomics researchers



For more details on zenvisage' novel design and how it works at scale, see our paper!

One Key Takeaway: Recommendations Kickstart Queries

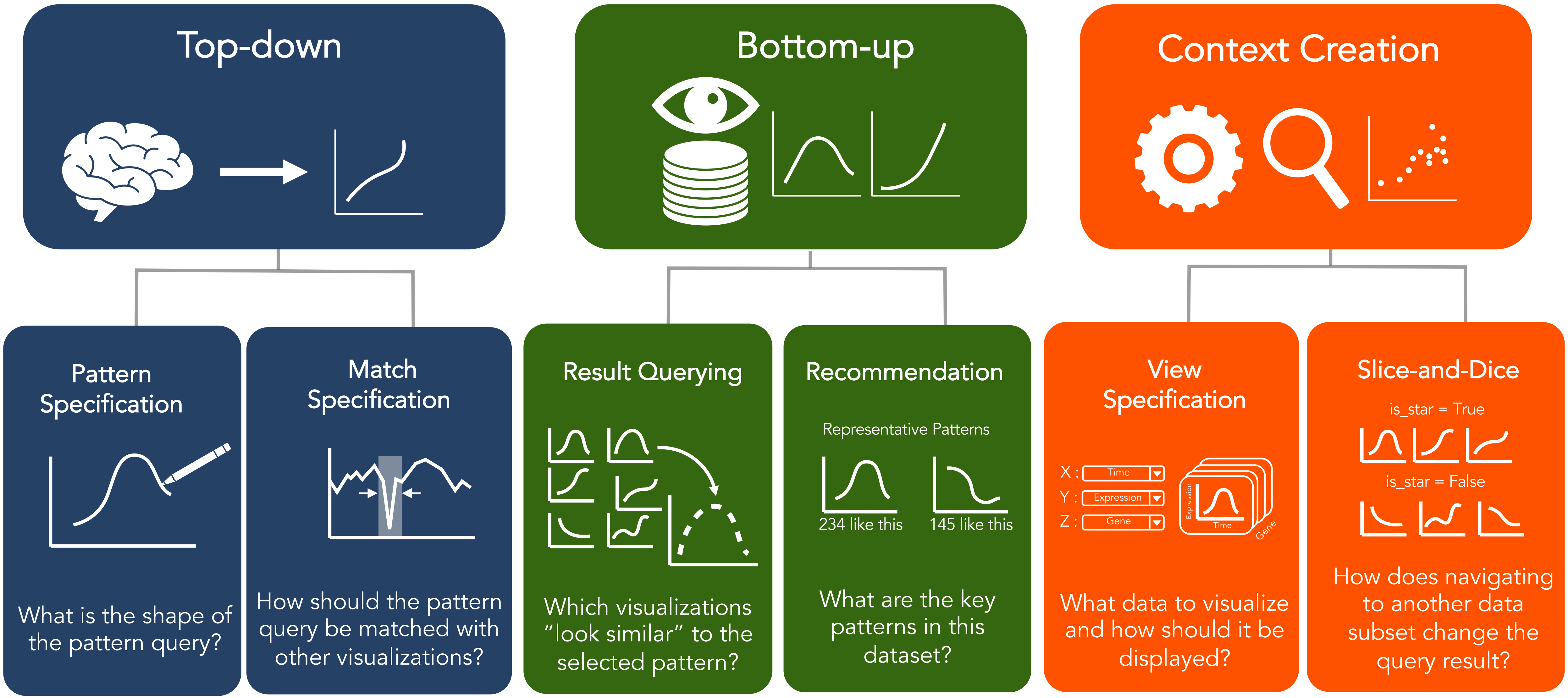


Direct sketch **only useful to 20%** of users

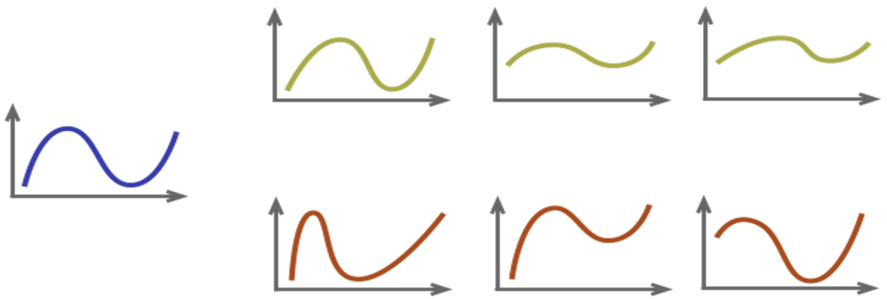
Drag & drop ➔ tweak preferred

Multiple input modalities help + Fostering a dialog helps

Taxonomy of Sensemaking Processes in VQSs

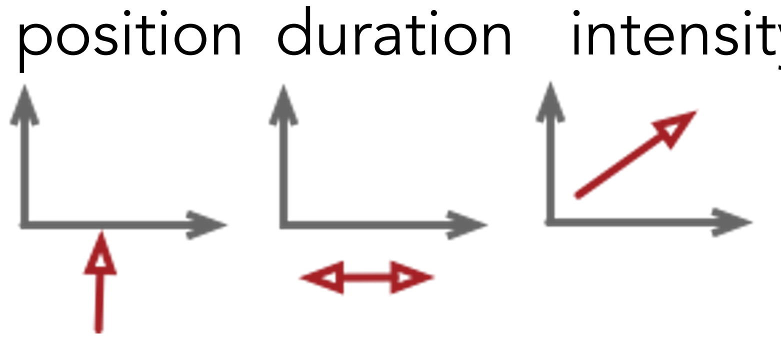


Unaddressed Needs in Visual Query Systems



Approximate search

“I just want to match the shape, not the values"



Partial (ambiguous) specification

“I don’t want to provide a precise specification – just key features"

AND

AT LEAST

CONCAT

OR

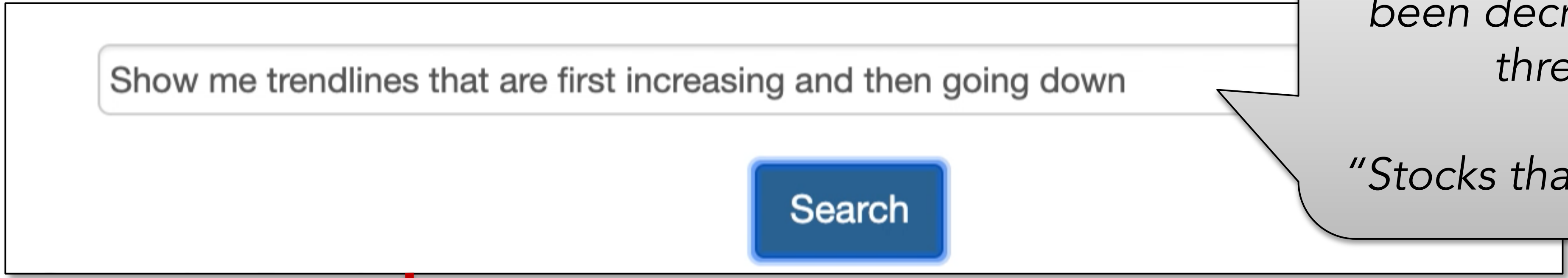
NOT

BETWEEN

Arbitrary combination of patterns

“I want to be able to mix and match characteristics”

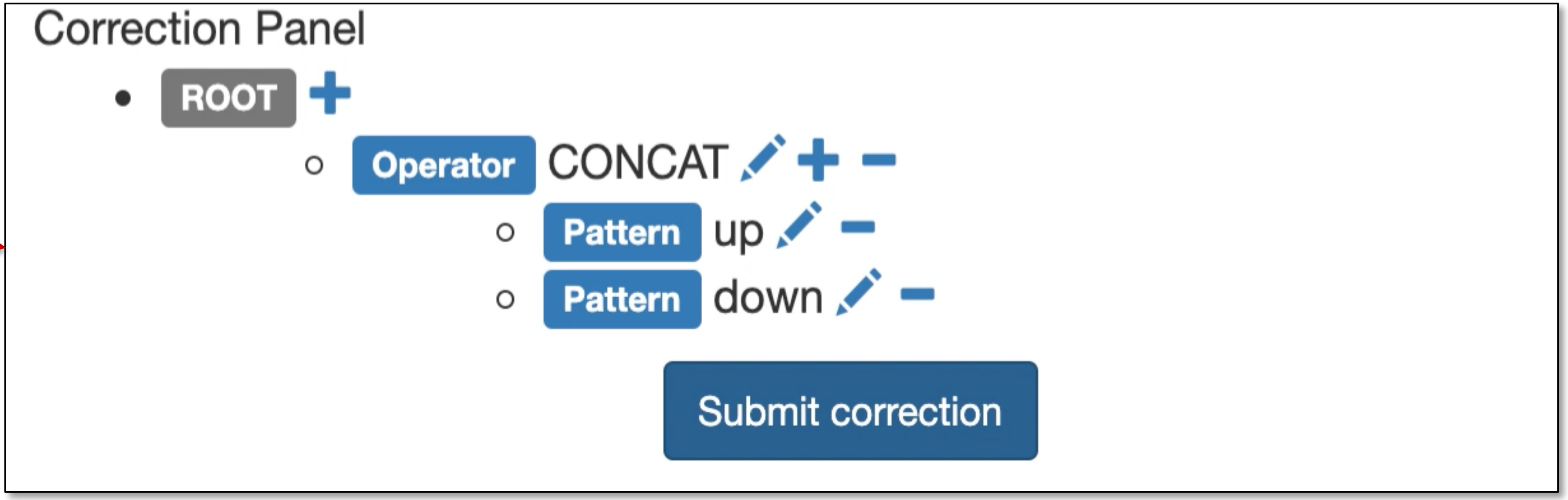
ShapeSearch: The Next Generation VQS



"Products whose sales have been decreasing in the last three months"

"Stocks that peaked in 2019"

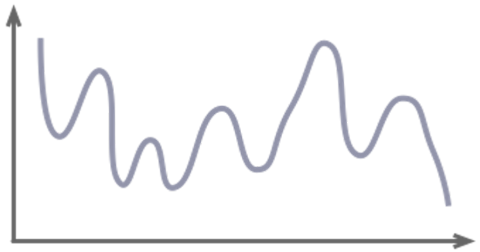
NLP Magic
[see paper!]

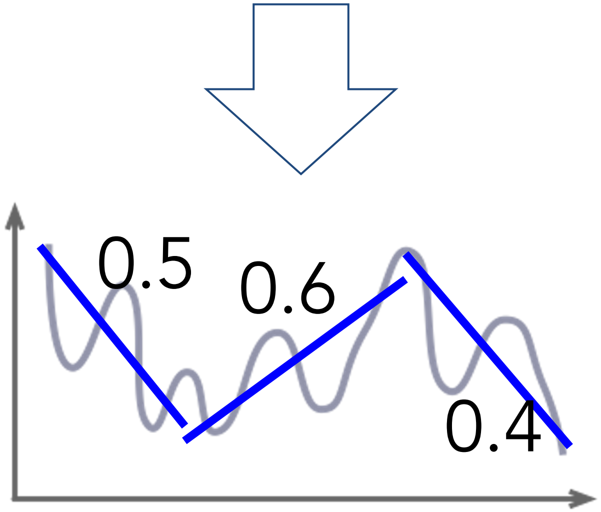


Keyword search for trendlines, with a panel showing the underlying system interpretation for validation (regex)

Fun with Algorithms: Scoring Patterns

Decreasing, then increasing and then decreasing +

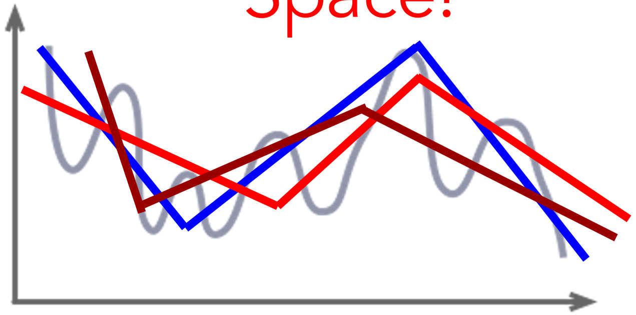




1. Approximate with lines
2. Score each pattern independently
3. Aggregate

$(0.5 + 0.6 + 0.4) / 3 = 0.5$

Challenge: Large Exploration Space!



$N^{(K-1)}$ possible approximations

(N = trend line length, K = no. of patterns)

Bottom-up algorithm: see paper!

Study Findings

No one modality dominates!

Keyword search dominates for:

- sequence-based patterns
- sub-patterns
- multiple constraints

➔ Supporting imprecision helps!

Sketching dominates for:

- precise pattern matching
- complex patterns

Case Study: Revisiting the Key Design Principles

1. Imprecision tolerated
   – sketch + keyword search offered valuable no-code ways to query vis

2. Human in a dialog with automation
   – important to support constructive dialog
   – kickstart with recommendations
   – confirm/validate when unclear
   – support multiple user workflows (three processes)

3. Multiple specification modalities
   – sketch, drag-and-drop, keyword search, …

See Dixin's talk later today!

Now to simply apply it to the rest of data work!

How do we make

training an ML model
building a dashboard
identifying patterns
data cleaning
…

as easy as **search**?

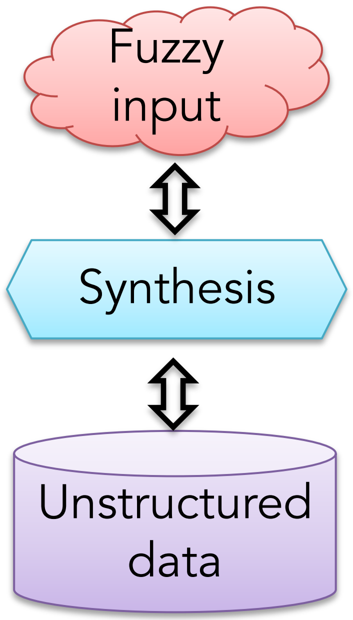
1. Imprecision tolerated
2. Human in a dialog with automation
3. Multiple specification modalities

Why Now?

Thanks to large pretrained models, we can now better

- Interpret imprecise input
- Synthesize data work programs automatically
- Understand unstructured data (PDFs, images, videos, …)

But it's still challenging to use these in practice!



	foo	bar	baz	zoo
0	one	A	1	7
1	one	B	2	8
2	one	C	3	9
3	two	A	4	10
4	two	B	5	11
5	two	C	6	12

```
# replace all cells that have 2 with 10
df.replace(2, 10)
```

```
# replace all cells that have 2 with 10 and all cells that have two with ten
df.replace({'foo': {'one': 'ten', 'two': 'twenty'}, 'baz': {2: 10}}, inplace=True)
```

No way for users to verify/guide/correct system output

See Sam and Justin's talks later today!

Why Now (2)

Impressive performance on understanding unstructured data

However, mistakes still abound — and are hard to fix.

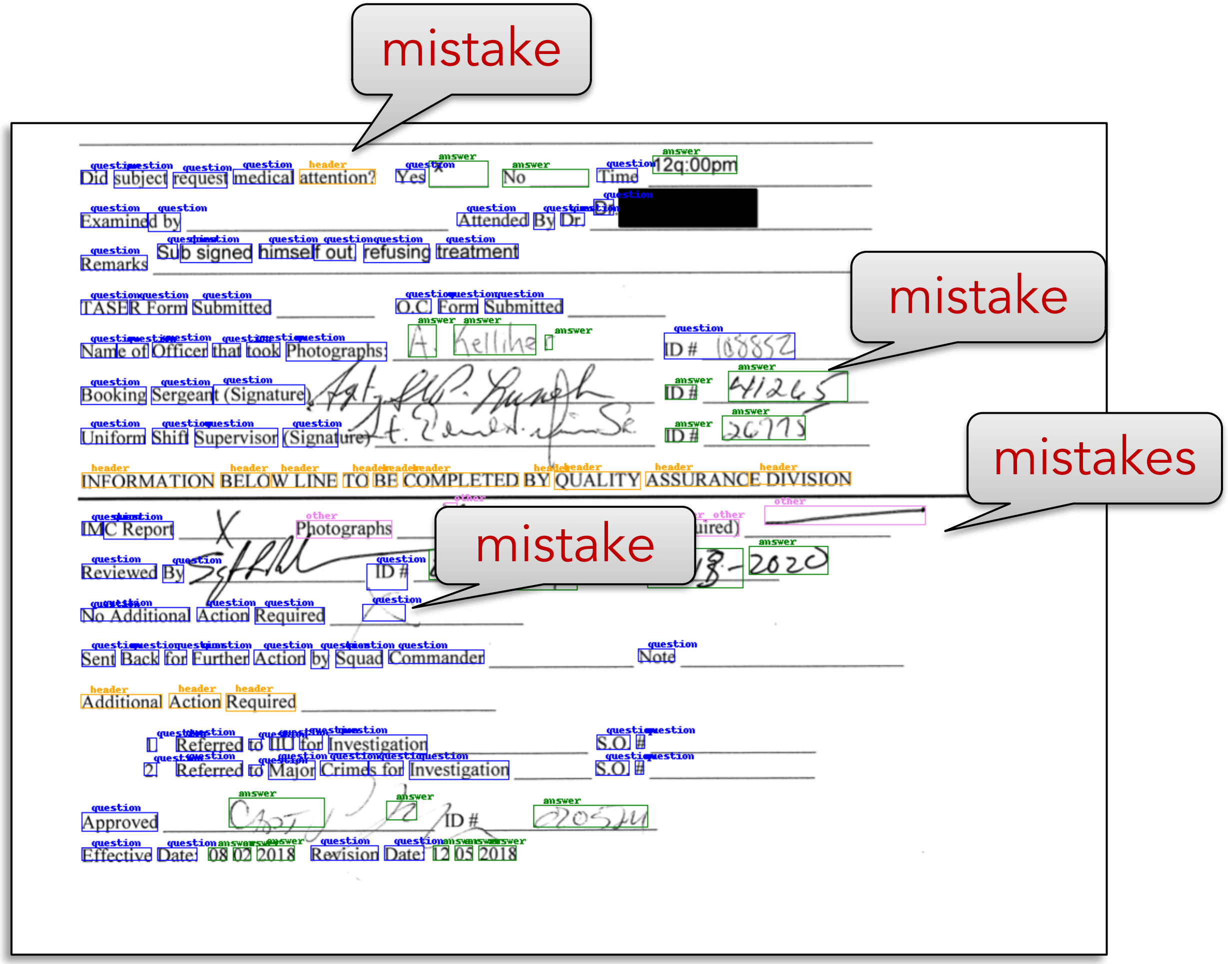


Image courtesy Manish Shetty
See Hellina and Rachel's talks later today!

Why Now? (3)

Playground

Load a preset... Save

Extract names of people from this text:

"Rishi Sunak (/ˈrɪʃɪ ˈsuːnæk/;[1] born 12 May 1980)[2] is a British politician who has served as Prime Minister of the United Kingdom and Leader of the Conservative Party since 25 October 2022. He was Chancellor of the Exchequer from 2020 to 2022 and Chief Secretary to the Treasury from 2019 to 2020.[3] He has been a member of Parliament (MP) for Richmond (Yorks) since 2015. Sunak was born in Southampton to parents of Indian descent who migrated to Britain from East Africa in the 1960s.[4][5][6] He was educated at Winchester College, earned a degree in philosophy, politics and economics (PPE) at Lincoln College, Oxford, and an MBA from Stanford University in California as a Fulbright Scholar. While at Stanford, he met his future wife Akshata Murty, the daughter of N. R. Narayana Murthy, the Indian billionaire businessman and co-founder of Infosys. Sunak and Murty are the 222nd richest people in Britain, with a combined fortune of £730m as of 2022.[7] After graduating, Sunak worked for Goldman Sachs and later as a partner at the hedge fund firms the Children's Investment Fund Management and Theleme Partners. His grandfathers were born in Punjab province, British India, and migrated from East Africa with their families to the UK in the 1960s.[19] His paternal grandfather, Ramdas Sunak, was from Gujranwala (present-day Pakistan) and moved to Nairobi in 1935 to work as a clerk, where he was joined by his wife Suhag Rani Sunak from Delhi in 1937.[20][21] His maternal grandfather, Raghubir Sain Berry MBE, worked in Tanganyika as a tax official, and had an arranged $\boxed{\text{marriage with 16-year-old Tanganyika-born Sraksha,}}$ with whom he had three children, and the family moved to the UK in 1966, funded by Sraksha selling her wedding jewellery.[22] In the UK, Raghubir Berry joined the Inland Revenue, and as a collector, was appointed a Member of the Order of the British Empire (MBE) in the 1988 Birthday Honours list.[23][24][25][26]"

Answer:

Rishi Sunak, Akshata Murty, N. R. Narayana Murthy, Ramdas Sunak, Suhag Rani Sunak, Raghubir Sain Berry

Submit 505

See Shreya talk later today!

Why Now? (4)

If we are to make insight discovery as easy as search, such design principles are key, while meeting interactivity requirements at scale.

1. Imprecision tolerated
2. Human in a dialog with automation
3. Multiple specification modalities