

Think outside the search box

Dr. Tony Russell-Rose FBCS CITP CEng



Professional search

- We all need search but...
 - Some need it more than others
- Patent, healthcare, recruitment, media...
 - Keyword search is not good enough
- Professional search needs to be:
 - Comprehensive
 - Repeatable
 - Transparent



Query builders

**World Health Organization**

**International Clinical Trials Registry Platform**
Search Portal

HomeAdvanced SearchList BySearch TipsUTNICTRP websiteREGTRACContact us

Fields can be left blank. Click on the field name hyperlink for an explanation of each search field

Look for trials with the exact phrase or contains

Example: liver cancer OR breast cancer

ANDExample: Diphtheria NOT tetanus

ANDExample: transplant AND immunosuppressant

in the [Title](#)

in the [Condition](#)

in the [Intervention](#)

☐ Without synonyms

☐ Without synonyms

☐ Search for [clinical trials in children](#)

[Recruitment status](#) is

Recruiting

[Primary sponsor](#) is or contains

[Secondary ID](#) is or contains

[Countries of recruitment](#) are

Afghanistan

Albania

Algeria

American Samoa

Andorra

Angola

Antigua and Barbuda

Argentina

Free Text Country :

>>

<<

Clear

[Date of registration](#) is between

dd/mm/yyyy

and

dd/mm/yyyy

[Phases](#) are

All

Phase 0

Phase 1

Phase 2

Phase 3

Phase 4

☐ With results only

☐ Rare diseases / orphan drugs

☐ Genome editing

Search

Clear

[Search Tips](#)

Disclaimer: Trials posted on this search portal are not endorsed by WHO, but are provided as a service to our users. In no event shall the World Health Organization be liable for any damages arising from the use of the information linked to in this section. None of the information obtained through use of the search portal should in any way be used in clinical care without consulting a physician or licensed health professional. WHO is not responsible for the accuracy, completeness and/or use made of the content displayed for any trial record.

[Copyright - World Health Organization](#) - Version 3.6 - [Version history](#)

3

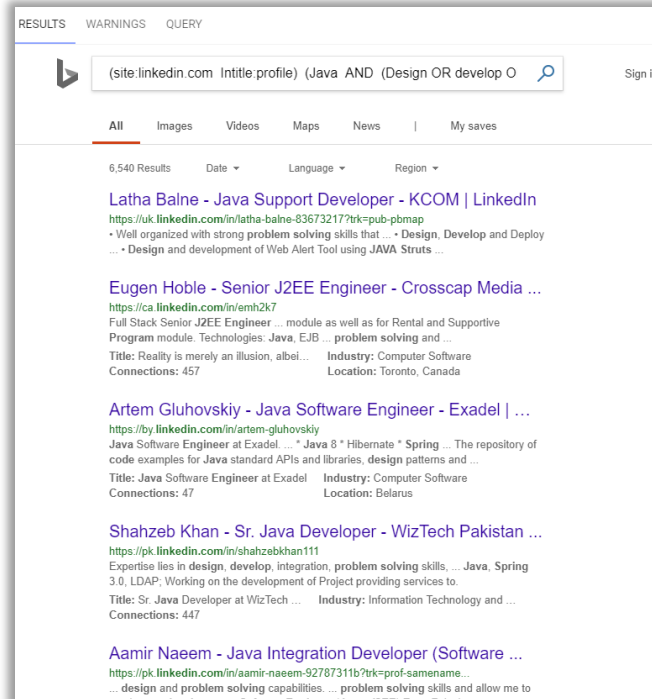
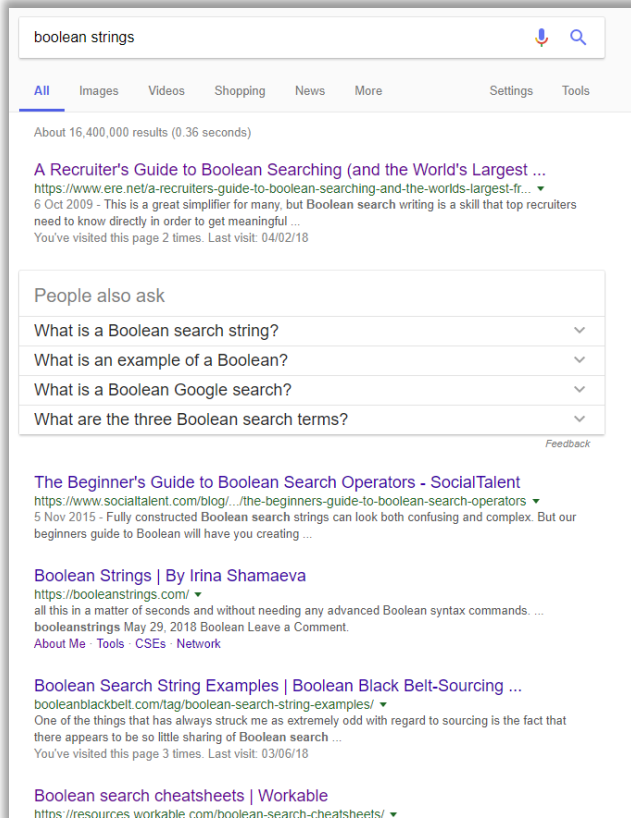
A healthcare search strategy

1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. randomized.ab.
4. placebo.ab.
5. clinical trials as topic.sh.
6. randomly.ab.
7. trial.ti.
8. 1 or 2 or 3 or 4 or 5 or 6 or 7
9. (animals not (humans and animals)).sh.
10. 8 not 9
11. exp Child/
12. ADOLESCENT/
13. exp infant/
14. child hospitalized/
15. adolescent hospitalized/
16. (child\$ or infant\$ or toddler\$ or adolescens\$ or teenage\$).tw.
17. or/11-16
18. Child Nutrition Sciences/
19. exp Dietary Proteins/
20. Dietary Supplements/
21. Dietetics/
22. or/18-21
23. exp Infant, Newborn/
24. exp Overweight/
25. exp Eating Disorders/
26. Athletes/
27. exp Sports/
28. exp Pregnancy/
29. exp Viruses/
30. (newborn\$ or obes\$ or "eating disorder\$" or pregnan\$ or childbirth or virus\$ or influenza).tw.
31. or/23-30
32. 10 and 17 and 22
33. 32 not 31

Francis, D. K., Smith, J., Saljuqi, T., & Watling, R. M. (2015). Oral protein calorie supplementation for children with chronic disease. *Cochrane Database of Systematic Reviews*, (5).

Recruitment search strategies

Java AND (Design OR develop OR code OR Program) AND ("* Engineer" OR MTS OR "* Develop*" OR Scientist OR technologist) AND (J2EE OR Struts OR Spring) AND (Algorithm OR "Data Structure" OR PS OR "Problem Solving")



Think outside the search box – www.2dsearch.com

Patent search & media monitoring

```
1      A01N0025-004/CPC
2      RODENT OR RAT OR RATS OR
MOUSE OR MICE
3      BAIT OR POISON
4      2 AND 3
5      1 OR 4
6      AVERSIVE OR ADVERSIVE OR
DETER? OR REPEL?
7      NONTARGET OR (NON WITH
TARGET) OR HUMAN OR DOMESTIC OR
PET OR DOG OR CAT
8      6 AND 7
9      8 AND 5
10     BITREX OR DENATONIUM OR
BITREXENE OR BITTERANT OR BITTER
11     10 AND 5
12     9 OR 11
```

```
(
  (
    (~Magnolia OR ~Magnolias) AND
    ('~Paul ~Anderson' OR '~Paul ~Thomas' OR paul* OR thom*
OR anders*)
  ) OR
  'jeremy blackman*' OR jeremy OR Jerem* OR
  (
    (~incredible OR ~coincidences) AND
    (role* OR chance* OR life* OR investigat* OR disturb*)
  ) OR
  (
    (~C NOT
      ('~John C' OR
        (
          (
            ('~Mister C' AND ~Shamen) OR
            ('~Tom C' AND ~Scientology)
          )
        )
      )
    )
  ) OR ~XCs OR ~CV OR ~CVs OR
  (
    (
      (~CKR OR ~CKRs) NOT ('CKR 123' OR lawyer)
    )
  )
)
```

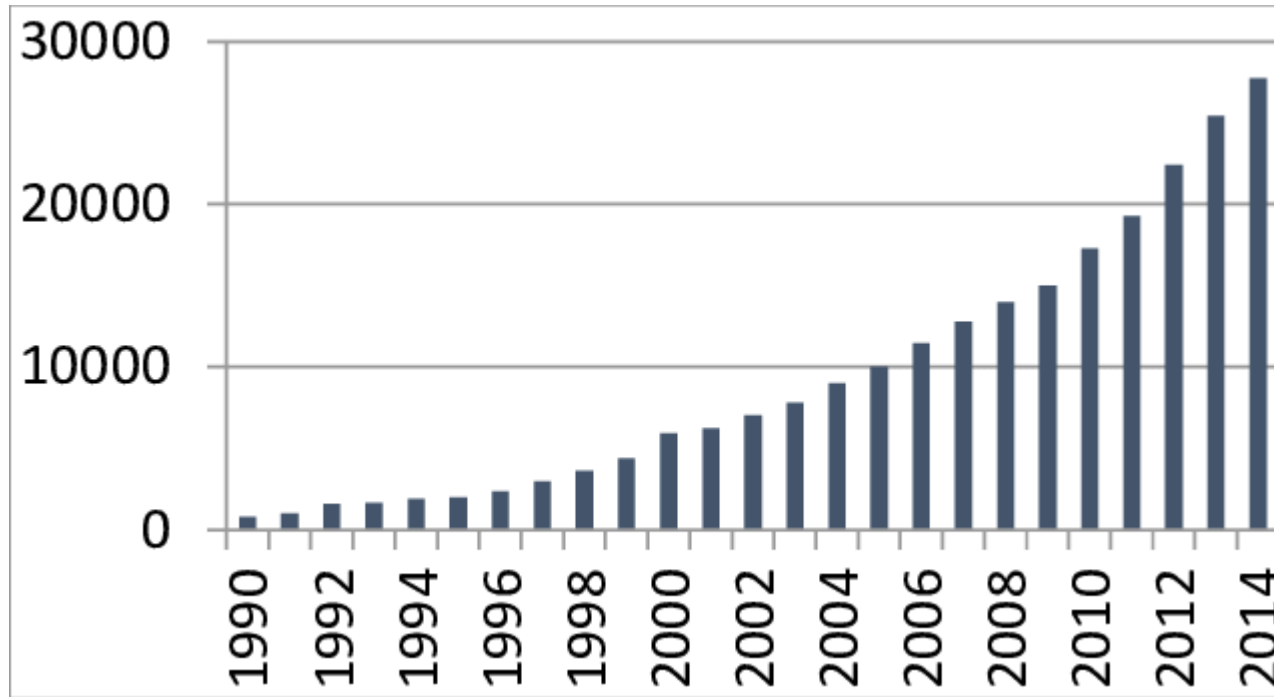
And more...

(topic:"d5f485b9\-aabo\-4d95\-a83b\-28b2decdd900c" AND (topic:"ba942587\-854e\-407d\-81dc\-812ee29624bd" OR topic:"805a2281\-8fo4\-42b7\-a451\-e2b7daac4db1" OR topic:"1fc2a4f5\-3f23\-4c80\-943c\-42f789d1644" OR topic:"91e4df78\-db55\-4d63\-be65\-2fad1afa7a7" OR organization:"Argos" OR organization:"Sainsbury" OR organization:"Marks and Spencer" OR organization:"Morrisons" OR organization:"Signet" OR organization:"Rapha" OR organization:"Vodafone" OR organization:"Evans Cycle" OR organization:"The Body Shop" OR organization:"Homebase" OR organization:"Topps Tiles" OR organization:"Carphone Warehouse" OR organization:"B&Q" OR organization:"Tesco" OR organization:"Lidl" OR organization:"Aldi" OR organization:"Waitrose" OR organization:"John Lewis" OR organization:"Dixons" OR organization:"Alliance Boots" OR organization:"AS Watson" OR organization:"Halfords" OR organization:"Vision Express" OR organization:"Pets at Home" OR organization:"Asda" OR organization:"Pret a Manger" OR organization:"Greggs" OR organization:"Majestic Wine" OR organization:"Ocado" OR organization:"Selfridges" OR organization:"Nike" OR organization:"Adidas" OR organization:"Walmart" OR organization:"Woolworths" OR organization:"Debenhams" OR organization:"Halfords" OR organization:"Punch Taverns" OR organization:"Monarch Group" OR organization:"Whitbread" OR organization:"Carnival Cruise" OR organization:"KFC" OR organization:"Yum!" OR organization:"Costa Coffee" OR organization:"Pizza Hut" OR organization:"Greene King" OR organization:"Carluccio's" OR organization:"Gondola" OR organization:"Starbucks" OR organization:"Sodexo" OR organization:"SSP" OR organization:"Nando's" OR organization:"Avis" OR organization:"Paddy Power" OR organization:"Easyjet" OR organization:"Emaar" OR organization:"Nuffield Health" OR organization:"Safestore" OR organization:"TUI" OR organization:"Merlin Entertainment" OR organization:"Virgin Atlantic" OR organization:"IHG" OR organization:"Virgin Active" OR organization:"Thomas Cook" OR organization:"Odeon Hilton" OR organization:"Pizza Express" OR organization:"Zizzi's" OR organization:"Gemfields" OR organization:"Orlebar Brown" OR organization:"Rebecca Trowell" OR organization:"Spring Studios" OR organization:"Loewe" OR organization:"Jimmy Choo" OR organization:"Burberry" OR organization:"Conde Nast" OR organization:"Zadig & Voltaire" OR organization:"Prada" OR organization:"Fendi" OR organization:"French Connection" OR organization:"Lyle & Scott" OR organization:"LK Bennett" OR organization:"Alexander Wang" OR organization:"Sandro" OR organization:"Isabel Marant" OR organization:"Rag & Bone" OR organization:"J Crew" OR organization:"Maje" OR organization:"Ralph Lauren" OR organization:"Tommy Hilfiger" OR organization:"Michael Kors" OR organization:"Kenzo" OR organization:"Pringle of Scotland" OR organization:"Heidi Klein" OR organization:"Louis Vuitton" OR organization:"Lanvin" OR organization:"Loewe" OR organization:"Chanel" OR organization:"Valentino" OR organization:"Celine" OR organization:"Dior" OR organization:"Alexander McQueen" OR organization:"Victoria Beckham" OR organization:"Sonia Rykel" OR organization:"Bally" OR organization:"Notonthehighstreet" OR organization:"Rupert Sanderson" OR organization:"Dunhill" OR organization:"De Beers" OR organization:"Swarovski" OR organization:"Hermes" OR organization:"Cartier" OR organization:"David Yurman" OR organization:"Mulberry" OR organization:"Tiffany & Co" OR organization:"Boucheron" OR organization:"Lancel" OR organization:"Marc Jacobs" OR organization:"Longchamp" OR organization:"Kate Spade" OR organization:"Salvatore Ferragamo" OR organization:"Asos" OR organization:"Net a porter" OR organization:"Astley Clarke" OR organization:"Fenwick" OR organization:"Yoox" OR organization:"Harrods" OR organization:"Selfridges" OR organization:"MyTheresa" OR organization:"Lane Crawford" OR organization:"Liberty of London" OR organization:"Gucci" OR organization:"Armani" OR organization:"Zara" OR organization:"Topshop" OR organization:"Arcadia" OR organization:"Balenciaga" OR organization:"Stella McCartney" OR organization:"Brioni" OR organization:"Bottega Veneta" OR organization:"Christopher Kane" OR organization:"Anya Hindmarch" OR organization:"Business of Fashion" OR organization:"3i" OR organization:"Active Private Equity" OR organization:"Advent International" OR organization:"Advent" OR organization:"Apollo Global Management" OR organization:"Apollo" OR organization:"Apax" OR organization:"Apax" OR organization:"Augmentum Capital" OR organization:"Bain" OR organization:"Endless" OR organization:"BC Partners" OR organization:"Better Capital" OR organization:"HuttonCollins" OR organization:"HuttonCollins" OR organization:"Bridgepoint" OR organization:"Business Growth Fund" OR organization:"CDR" OR organization:"Darwin" OR organization:"Duke Street Capital" OR organization:"Electra Partners" OR organization:"Exponent" OR organization:"European Capital" OR organization:"General Atlantic" OR organization:"Hilco" OR organization:"Inflexion" OR organization:"KKR" OR organization:"LGV" OR organization:"Lion" OR organization:"Magenta Partners" OR organization:"Montagu" OR organization:"Octopus" OR organization:"Octopus Capital" OR organization:"Piper Private Equity" OR organization:"Piper" OR organization:"Sun Capital" OR organization:"TA Associates" OR organization:"Warburg Pincus" OR organization:"Charterhouse" OR organization:"Primary Capital" OR organization:"Bowmark" OR organization:"Beringea" OR organization:"Business Growth Fund" OR organization:"Bridgepoint" OR organization:"Caird Capital" OR organization:"CapVest" OR organization:"Cerberus Capital" OR organization:"CBPE" OR organization:"CCMP" OR organization:"CDR" OR organization:"Change Capital" OR organization:"Cinven" OR organization:"CVC" OR organization:"Duke Street" OR organization:"EC1" OR organization:"Electra" OR organization:"Encore" OR organization:"EQT" OR organization:"Equistone" OR organization:"Graphite Capital" OR organization:"Caledonian" OR organization:"Kings Park Capital" OR organization:"L Capital" OR organization:"LDC"

Healthcare search

- Evidence based medicine
- *"A systematic review summarises the results of healthcare studies (controlled trials) and provides a high level of evidence on the effectiveness of healthcare interventions. Judgments may be made about the evidence and inform recommendations for healthcare."*
- Cochrane Consumer Network, <https://consumers.cochrane.org/what-systematic-review>

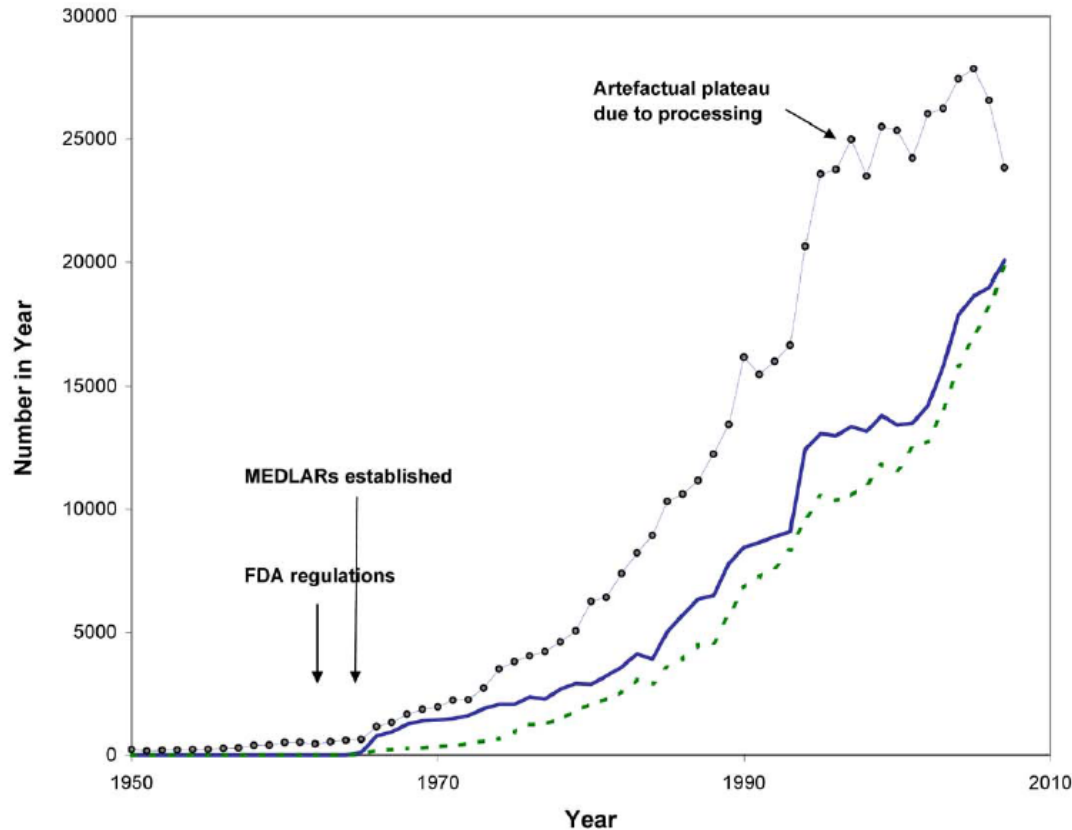
Big data in healthcare



Growth in systematic reviews over 25 years (N=233,785 MEDLINE citations)

Bekhuis T, Tseytlin E, Mitchell KJ. A Prototype for a Hybrid System to Support Systematic Review Teams: A Case Study of Organ Transplantation. *Proceedings IEEE International Conference on Bioinformatics and Biomedicine*. 2015;2015:940-947.

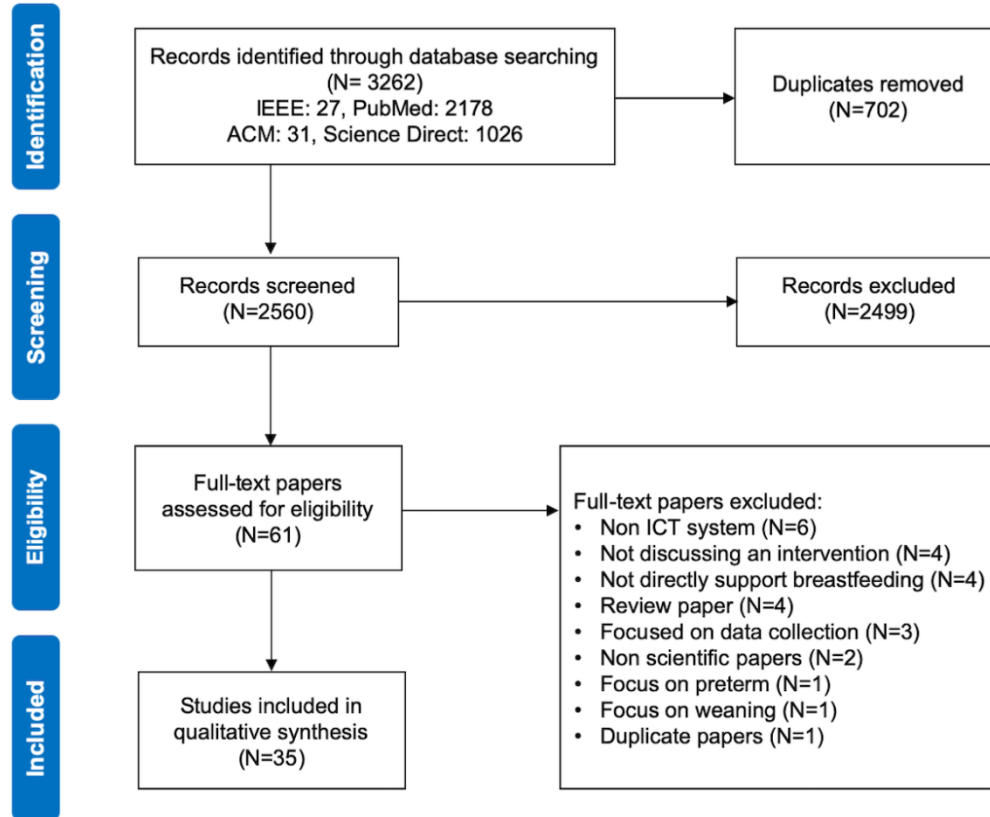
Big data in healthcare



The number of published trials,
1950 to 2007

Bastian H, Glasziou P, Chalmers I.
Seventy-five trials and eleven systematic
reviews a day: how will we ever keep up?
PLoS Med 2010 Sep 21;7(9)

PRISMA flowchart



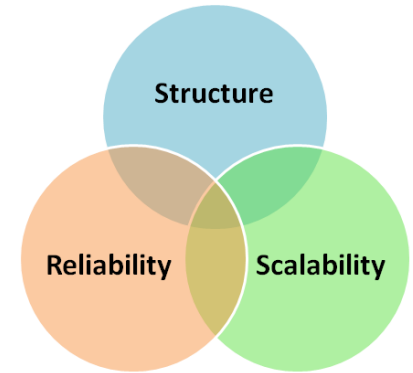
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) is an evidence-based minimum set of items aimed at helping authors to report a wide array of [systematic reviews](https://en.wikipedia.org/wiki/Systematic_Review) and [meta-analyses](https://en.wikipedia.org/wiki/Meta-analysis) that assess the benefits and harms of a health care intervention.

https://en.wikipedia.org/wiki/Preferred_Reporting_Items_for_Systematic_Reviews_and_Meta-Analyses

Tang, K., Gerling, K., Chen, W., & Geurts, L. (2019). Information and Communication Systems to Tackle Barriers to Breastfeeding: Systematic Search and Review. *Journal of Medical Internet Research*, 21(9), e13947.

What's wrong with Boolean strings?

- Ineffective for communicating *structure*
 - Visual cues needed
- Poor scalability
 - No abstraction
- Error-prone
 - Syntactic + semantic errors



- Salvador-Oliván, José Antonio, Gonzalo Marco-Cuenca, and Rosario Arquero-Avilés. "Errors in search strategies used in systematic reviews and their effects on information retrieval." *Journal of the Medical Library Association: JMLA* 107.2 (2019): 210.

An imperfect process

- Auditing published search strategies
 - Sampson M, McGowan J. Errors in search strategies were identified by type and frequency. J Clin Epidemiol 2006 Oct;59(10):1057-1063
 - Salvador-Oliván, J.A., Marco-Cuenca, G. and Arquero-Avilés, R., 2019. Errors in search strategies used in systematic reviews and their effects on information retrieval. Journal of the Medical Library Association: JMLA, 107(2), p.210.
- 63 MEDLINE search strategies examined
 - 90.5% contained ≥ 1 errors
 - 82.5% contained errors that could potentially lower recall of relevant studies
 - missed MeSH terms (44.4%)
 - unwarranted explosion of MeSH terms (38.1%)
 - irrelevant MeSH or free text terms (28.6%)
 - Missed spelling variants, combining MeSH and free text terms in the same line, and failure to tailor the search strategy for other databases (20.6%)
 - Logical operator error occurred in 19.0% of searches

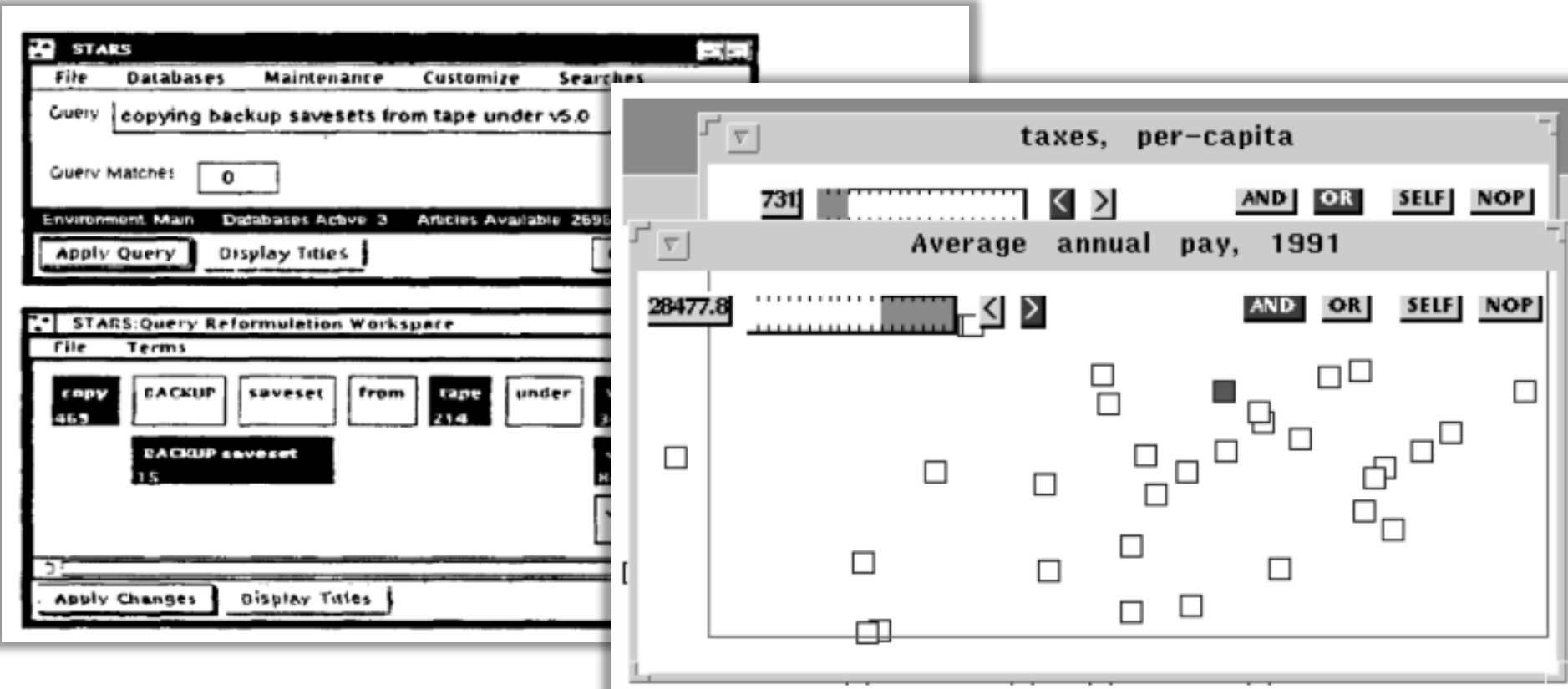
Learning from other disciplines

- Search strategies are:
 - Complex, logical constructs
 - Read many more times that they are written
 - Exhibit abstraction, modularity, encapsulation
 - Maintained, shared, optimised...
- *"For the IBM PC in the early 1980s, the software development process was text edit, compile, write down the errors, and debug with your eyes."*
 - The History of Visual Development Environments, <https://www.mendix.com/blog/the-history-of-visual-development-environments-imagine-theres-no-ides-its-difficult-if-you-try/>

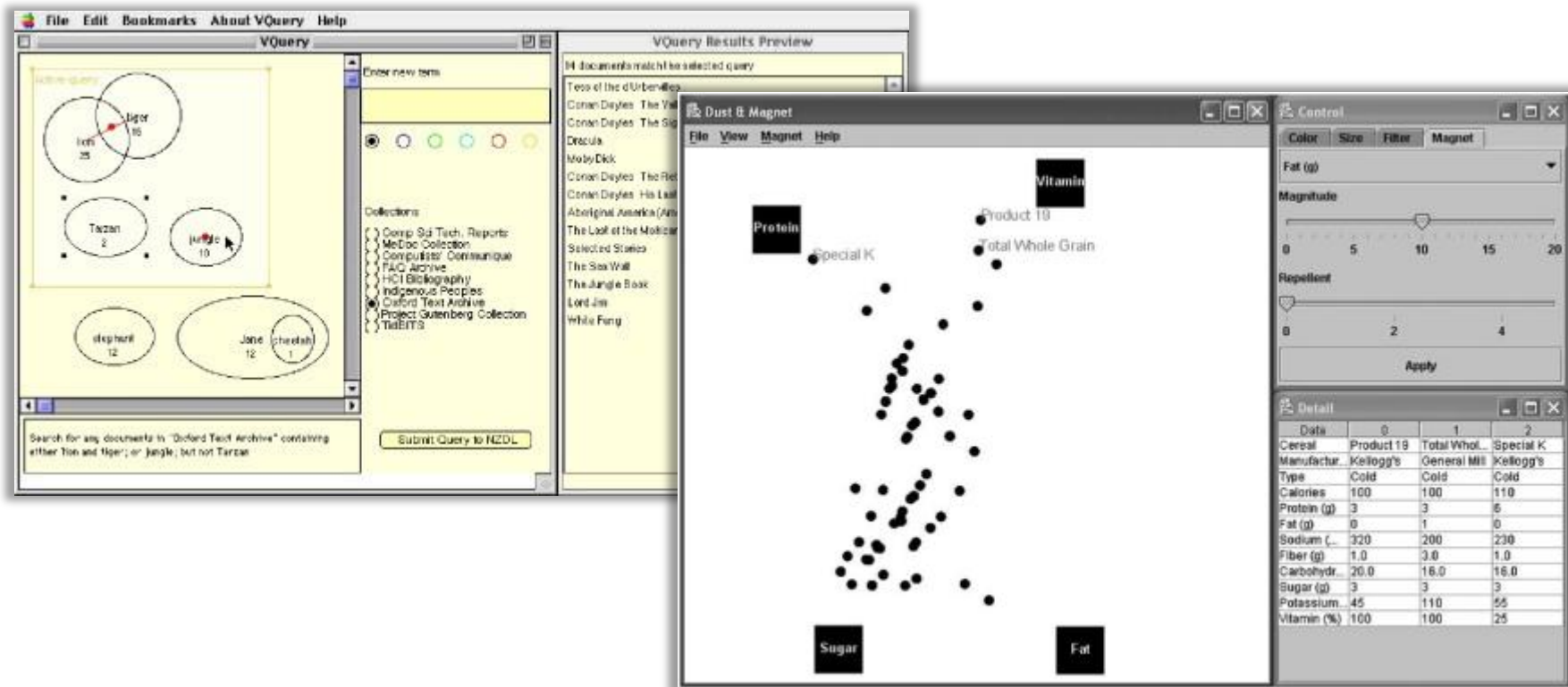
Thinking outside the search box

- What if we could *visualize* search strategies?
 - Use metaphors the user already understands
 - Separate keyword selection from query manipulation
 - Provide a 'scratchpad' & 'debugging aids'
 - Provide instant feedback
 - Support abstraction (levels of composition)
 - Provide query suggestions

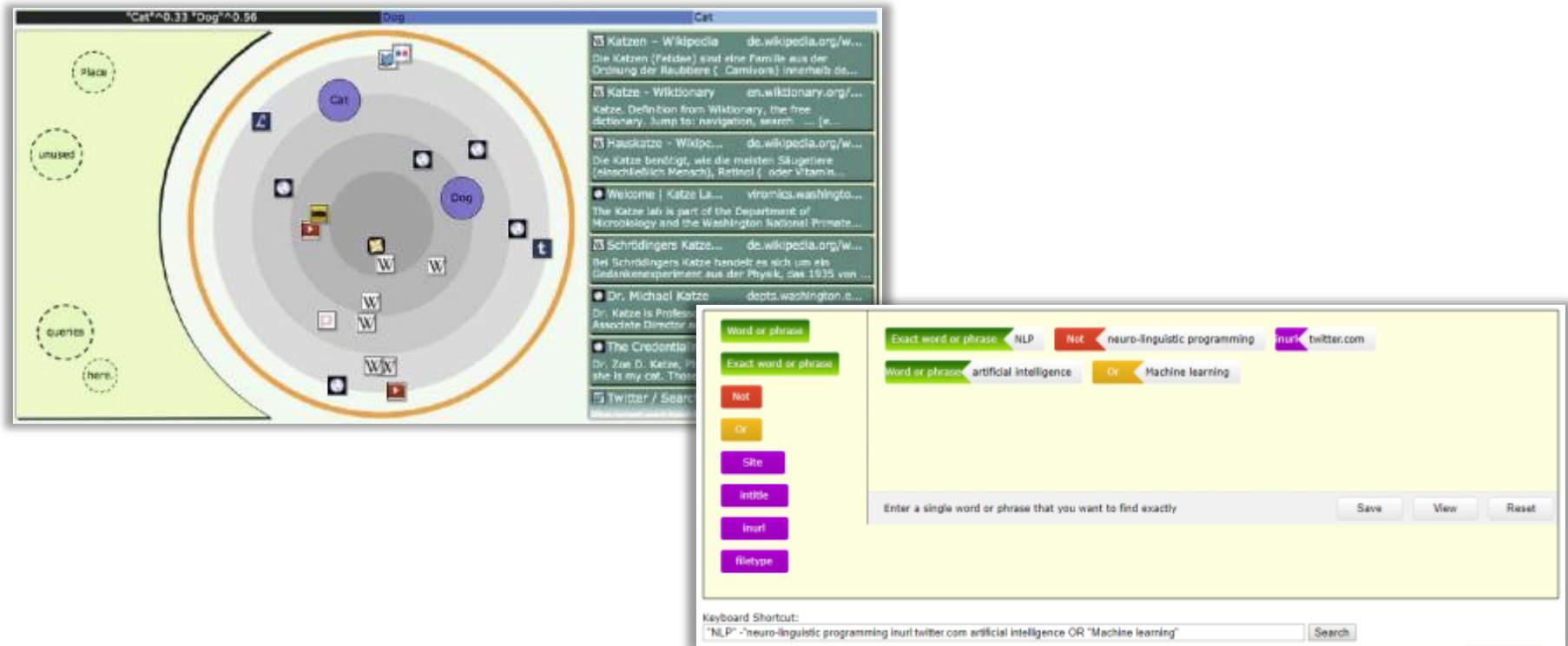
Anick et al (1989), Fishkin & Stone (1995)



Jones (1998), Yi et al (2005)



Nitsche & Nurnberger (2013), Boolify (2014?)



Boolio, Search Whiteboard, Searchrefiner

AND OR NOT RESET

Filetype - Off

LinkedIn

Software Enngineer	Platfon
Software Developer	Softwa
	PaaS
	SaaS

("Software Enngineer" OR "Software

Patient	Interve
Hypertension	aspirin
high blood pressure	NSAIDS
"Hypertension"	"aspirin"
"high blood pressure"	"NSAIDS"
Hypertension[TI]	aspirin[TI]
"high blood pressure"[TI]	NSAIDS[TI]
"high blood pressure"[OT]	aspirin[MeSH]
Hypertension[MeSH]	Anti-Inflammat Non-Steroidal[

searchrefiner

Update Tree

Text Editor Structured Editor Tree View Console

1. exp ORTHODONTICS/
2. orthodontic\$.mp.
3. or/1-2
4. (retention or retain\$).mp.
5. (stabilise\$ or stabilize\$).mp.
6. (fraenectom\$ or frenectom\$).mp.
7. (fiberotom\$ or fibreotom\$).mp.
8. interproximal stripping.mp.
9. pericision.mp.
10. reproxinat\$.mp.
11. ((gingiv\$ or periodont\$).mp. adj4
surg\$).mp.
12. (retain or retention).mp.
13. 11 and 12
14. or/4-10
15. 13 or 14
16. 3 and 15

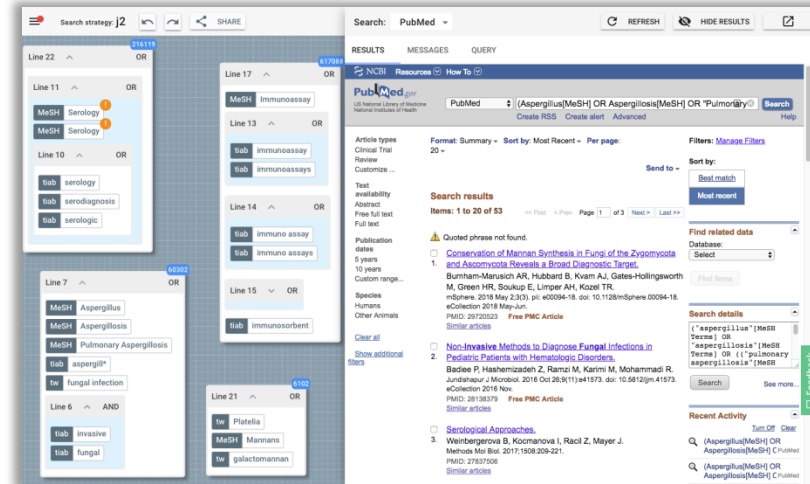
Query Language help?
Ovid MEDLINE

Results > Edit >

4161 citations retrieved
10 citations relevant
7 citations relevant retrieved
help?

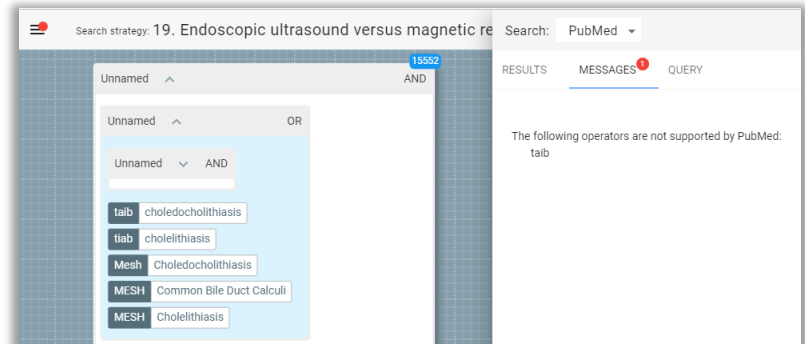
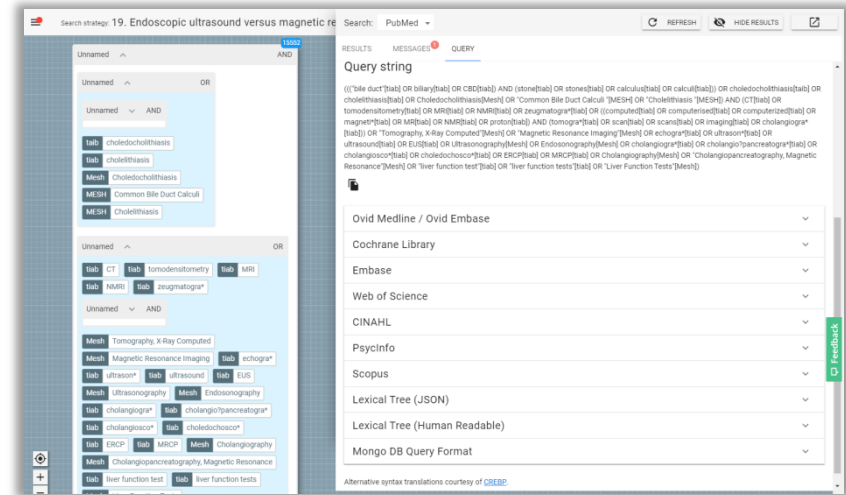
Query formulation

- Concepts are objects on a 2D canvas
- Relationships via direct manipulation
 - Syntactically correct expressions
 - Transparent semantics
 - Arbitrary levels of abstraction
- 'Lego' analogy
 - Optimisation thru exploration
 - Execute, enable/disable
 - Hit counts



Executing queries

- Integrations
 - 8 databases & counting...
- Universal representation
 - Auto-translation
 - Error-checking
- Interoperability with legacy formats
 - Boolean string I/O



That healthcare strategy...

1198210

Type of study ^ AND

RCT ^ OR

pt randomized controlled trial

pt controlled clinical trial

tiab randomized

tiab placebo

Mesh:NoExp clinical trials as topic

tiab randomly

ti trial

Humans only ^ AND

animals

Unnamed ^ AND

humans

animals

3994866

Target population ^ OR

Mesh Child

Mesh:NoExp ADOLESCENT

Mesh infant

Mesh:NoExp child hospitalized

Mesh:NoExp adolescent hospitalized

tiab child* tiab infant*

tiab toddler*

tiab adolescent*

tiab teenage*

161355

Intervention ^ OR

Mesh:NoExp Child Nutrition Sciences

Mesh Dietary Proteins

Mesh:NoExp Dietary Supplements

Mesh:NoExp Dietetics

3024636

Exclusions ^ OR

Mesh Infant, Newborn

Mesh Overweight

Mesh Eating Disorders

Mesh:NoExp Athletes

Mesh Sports

Mesh Pregnancy

Mesh Viruses tiab newborn*

tiab obes*

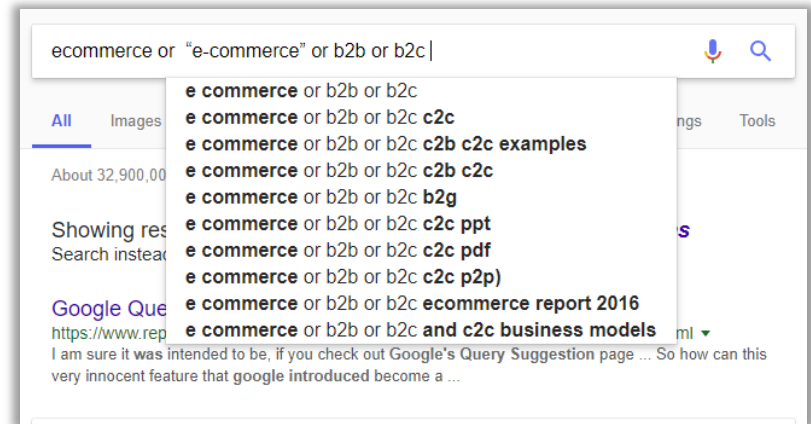
tiab eating disorder*

tiab pregnan* tiab childbirth

tiab virus* tiab influenza

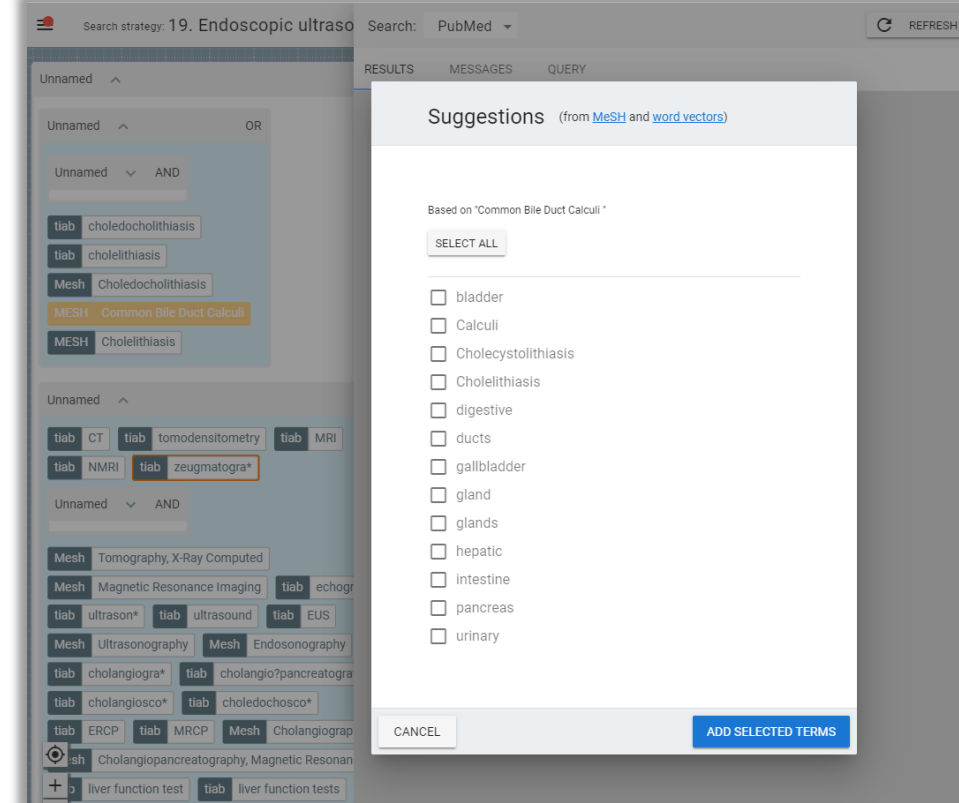
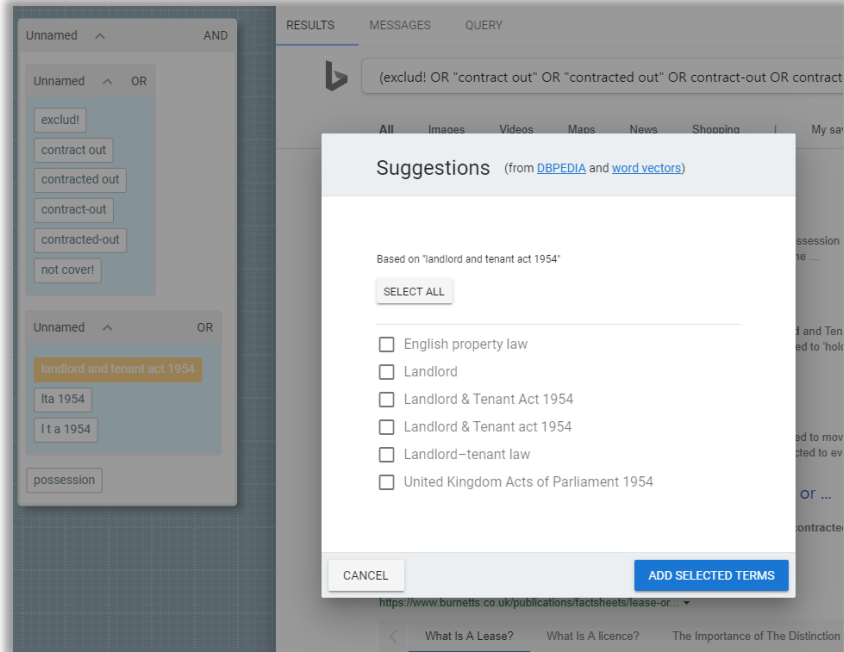
But what about content?

- Term selection is common source of inefficiency & error
 - Apply query suggestions?

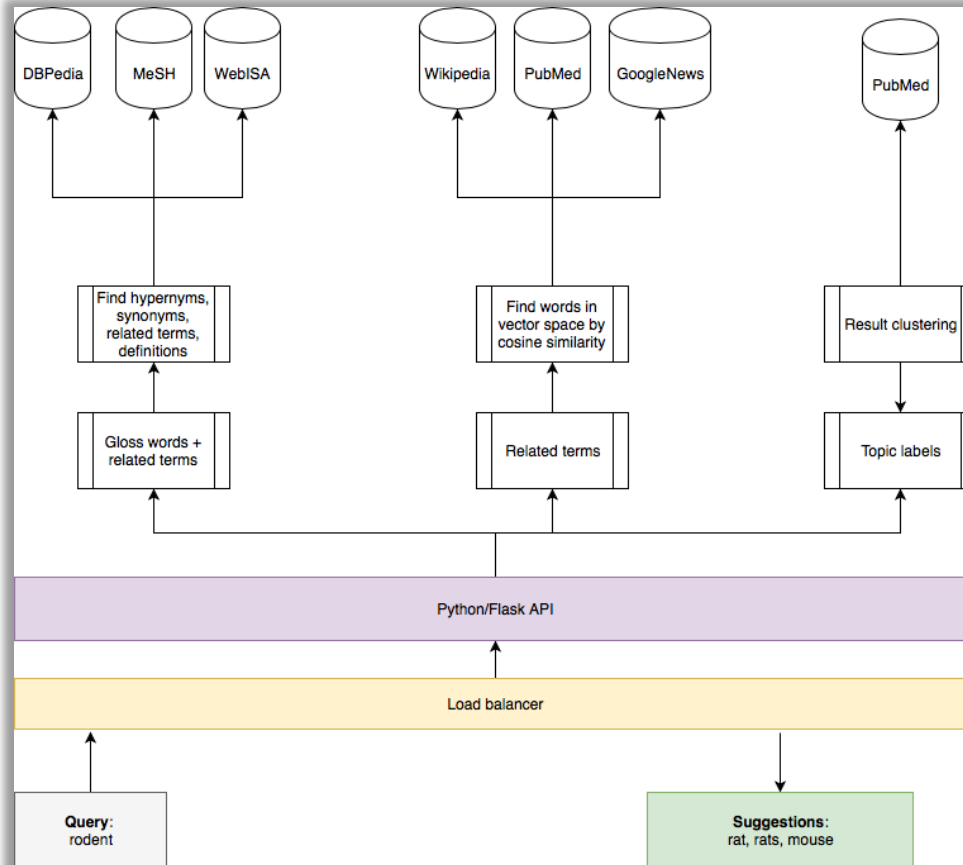


- (*"business analyst" or "systems analyst" or "system analyst" or "data analyst" or "requirements analyst" or "functional analyst"*) and *crystal* and *report** and *analy** and *data near analy** and not *inventory* and not *retail* and not (*ecommerce or "e-commerce" or b2b or b2c*)

Suggestions in context



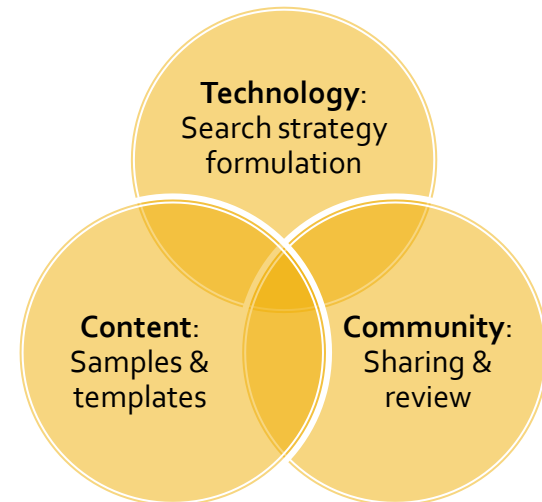
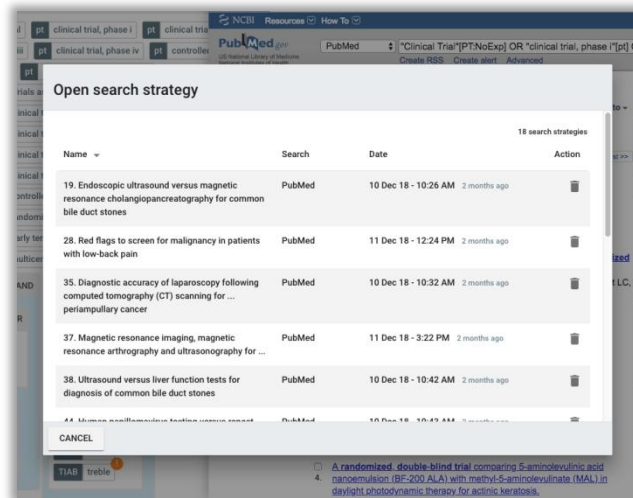
NLP architecture



T.G. Russell-Rose and Phil Gooch, "2dSearch: a visual approach to search strategy formulation", Proceedings of DESIRES: Design of Experimental Search & Information REtrieval Systems, Bertinoro, Italy, 28-31 August 2018

Sharing, reuse, best practice

- Open content repository:
 - Templates and best practices
 - Library of shared strategies (aka 'GitHub')
 - Community of practice (aka 'Stack Overflow')



Further work

- User experience
 - Visual methods vs. traditional query builders
 - Quality of query suggestions
- Data science
 - Semi-automated strategy generation
 - Investigation of ConceptNet, Wikidata, etc.
 - Combining set retrieval with ranked retrieval
 - Contextual language modelling (BERT, ELMo, etc.)

Summary

- Reimagining 'advanced search'
 - Eliminate syntactic error
 - Make semantics transparent
 - Provide support for reuse, optimisation, best practice
- Try it for yourself: <https://www.2dsearch.com>
 - Tutorial videos: <https://www.2dsearch.com/help>

Thank you

Dr. Tony Russell-Rose, FBCS CITP CEng

Senior Lecturer, Goldsmiths

Founder: 2Dsearch, Director: UX Labs

RAE Visiting Professor, Essex University

- Web: <http://www.2dsearch.com>
- Email: info@2dsearch.com
- Blog: <http://isquared.wordpress.com>
- LinkedIn: <http://uk.linkedin.com/in/tonyrussellrose>
- Twitter: @tonygrr, @2d_search