

Searching fast & slow

Dr. Tony Russell-Rose FBCS CITP CEng







Professional search

We all need search but...

- Google bing
- 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



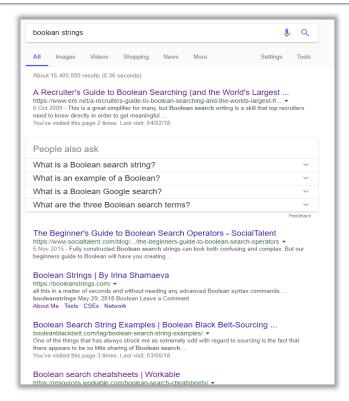
A healthcare search strategy

- 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. (childs or infants or toddlers or adolescens or teenages).tw.

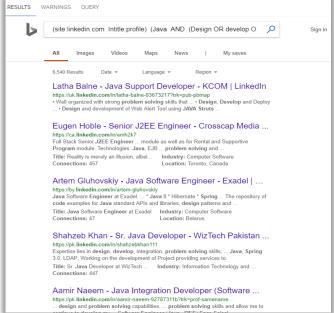
- 17. or/11-16
- 18. Child Nutrition Sciences/
- 19. exp Dietary Proteins/
- 20. Dietary Supplements/
- 21. Dietetics/
- 22. or/18-21
- 23. exp Infant, Newborn/
- 24. exp Overweight/
- 25. exp Eating Disorders/
- 26. Athletes/
- 27. exp Sports/
- 28. exp Pregnancy/
- 29. exp Viruses/
- 30. (newborns or obess or "eating disorders" or pregnans or childbirth or viruss 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")



Patent search & media monitoring

```
A01N0025-004/CPC
       RODENT OR RAT OR RATS OR
MOUSE OR MICE
       BATT OR POTSON
       2 AND 3
       1 OR 4
       AVERSIVE OR ADVERSIVE OR
       OR REPEL?
DETER?
       NONTARGET OR (NON WITH
        OR HUMAN
                 OR
                    DOMESTIC OR
       DOG OR CAT
       6 AND
         AND 5
       BITREX OR DENATORIUM
BITREXENE OR BITTERANT OR BITTER
11
       10 AND 5
       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)
```

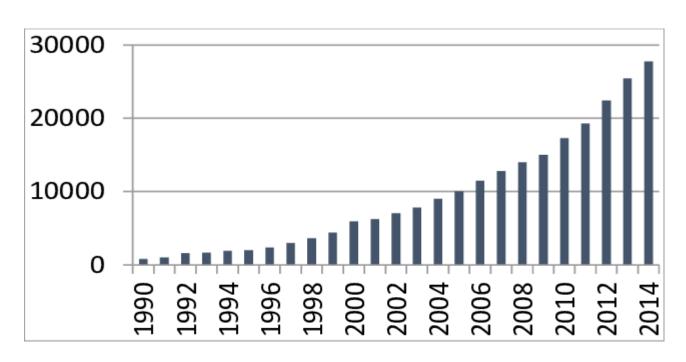
And more...

(topic:"dsf485b9\-aabo\-4d95\-a83b\-28b2decd9ooc" AND (topic:"ba942587\-854e\-407d\-81dc\-812ee29624bd" OR topic:"805a2281\-8f04\-42b7\-a451\-e2b7daac4db1" OR topic:"1fc2a4f5\-3f23\-4c80\-9436\-42f789d11644" OR topic: "91e4df78\-db55\-4d63\-be65\-2fad1afa77a7" 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: "Zadiq & 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: "Chanel" 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: "Armarni" 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:"ai" 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: "Apax" OR organization: "BC orga 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: "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" 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: "CapVest" 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: "Činven" OR organization: "ČVC" 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"
Think outside the search box — www.2dsearch.com

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, <u>https://consumers.cochrane.org/what-systematic-review</u>

Big data in healthcare

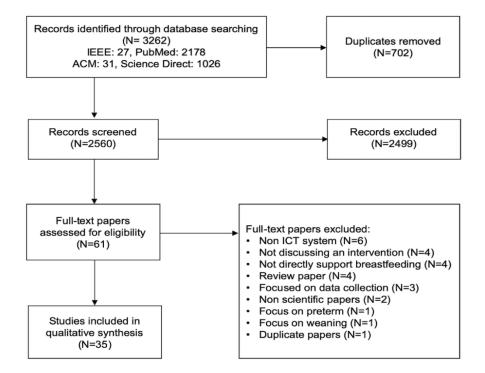


Growth in systematic reviews over 25 years (N=233,785 MEDLINE citations)

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.

PRISMA flowchart

Identification



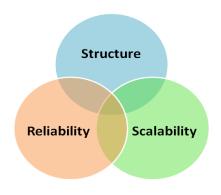
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) is an evidence-based minimum set of items aimed at helping authors to report a wide array of systematic reviews and meta-analyses that assess the benefits and harms of a health care intervention.

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.



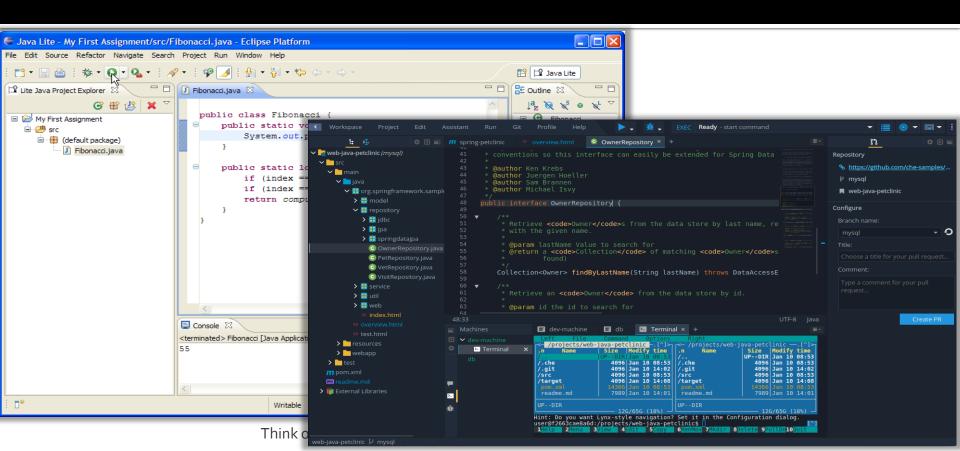
Professional search practice

- Peer-reviewing search strategies (McGowan et al. 2016, J Clin Epidemiology)
- Search strategy development (Russell-Rose & Shokraneh, Journal of Library UX)
 - Word Processors
 - Auto-conversion of straight to smart quotations
 - Auto-correction of truncation or wildcards: Creates duplicates
 - Auto-space-removal of shared documents
 - Need external program such as Microsoft Office or GoogleDocs
 - PDFs
 - Breaking paragraphs into lines when copies
 - Cannot be edited (without a installing another program)
 - Need external program such as Microsoft Office or GoogleDocs

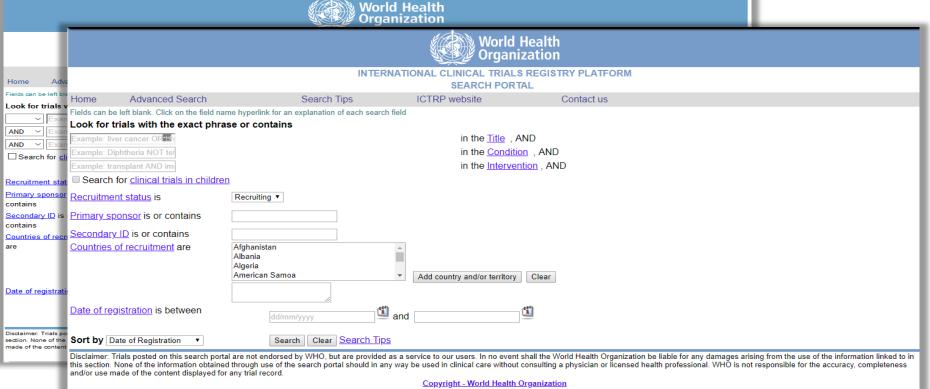
Professional search practice

- Reproducibility of searches
- Training beginners
- Translating search strategies
 - PolyGlot
- Sharing search strategies (externally)
 - Unavailability of supplementary materials (Evangelou et al. 2005, FACEB J)
 - MedTerm Assist (Pittsburgh University), Search Strategy Library (Gulhane, 2019)
 - ISSG Search Filters, Search Blocks on Google Sites
 - Institutional Repositories or OSF
- Living Search Strategies (Shokraneh & Russell-Rose 2020, J Clin Epidemiol)
 - COVID-19

Quiz time! 2010 or 2020?



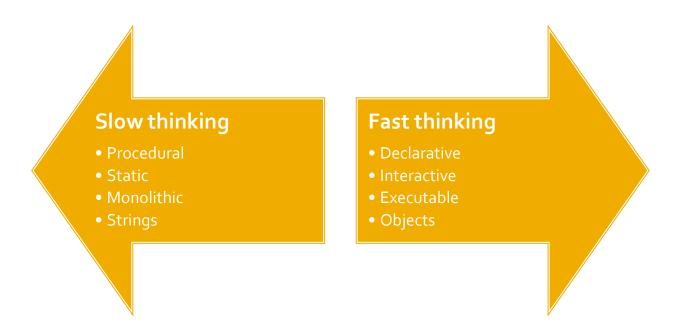
Quiz time! 2010 or 2020?



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/

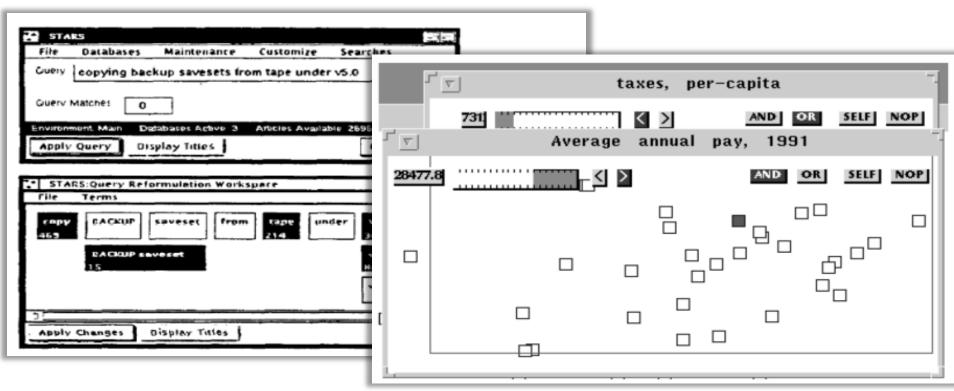
Searching fast & slow



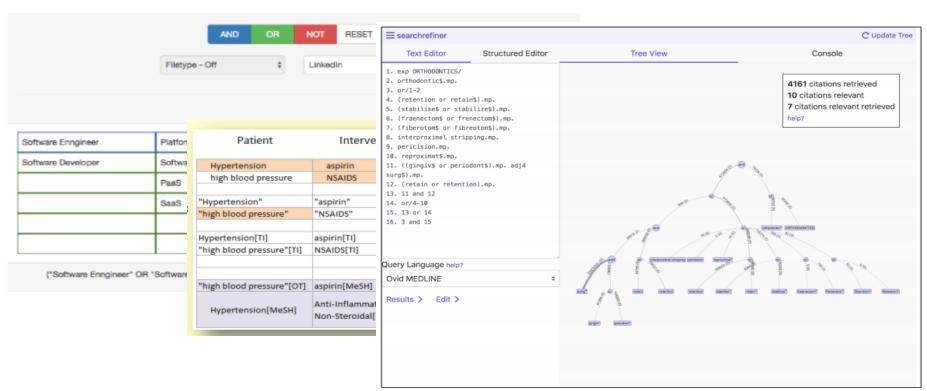
Thinking outside the search box

- 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)

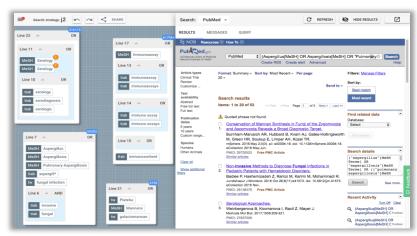


Boolio, Search Whiteboard, Searchrefiner



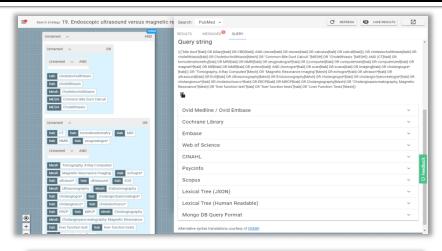
Query formulation

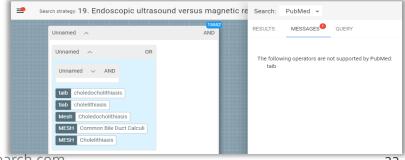
- Concepts are objects on a 2D canvas
- Relationships via direct manipulation
 - Syntactically correct expressions
 - Transparent semantics
 - Arbitrary levels of abstraction
- 'Lego' analogy
 - Optimisation thru exploration
 - Execute, enable/disable
 - Hit counts



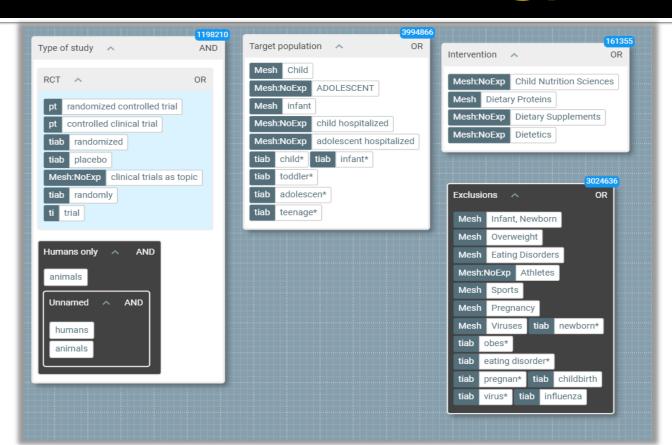
Executing queries

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

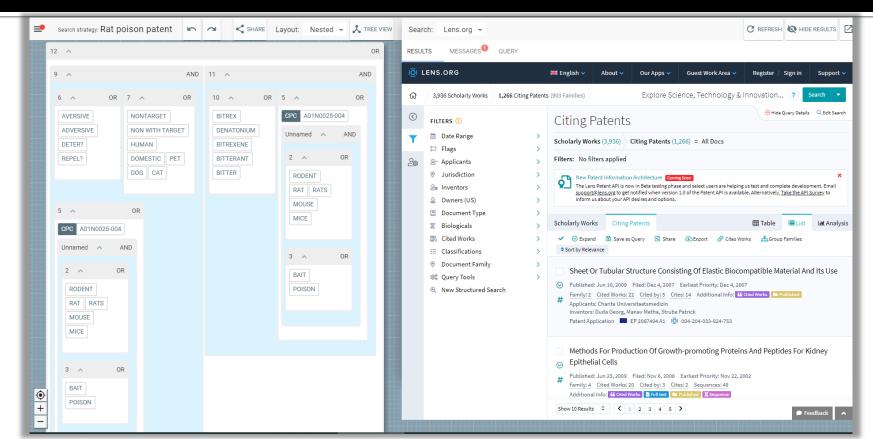




That healthcare strategy...



That patent strategy...



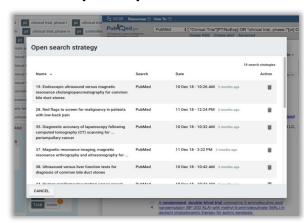
Design principles

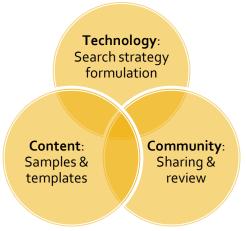
- 1. Support **transparency** in the **mapping** between logical structure and physical structure
- 2. Adopt **scalable formalisms** that accommodate complexity
- 3. Delegate lower-level syntactic operations to system functions
- 4. Provide **real-time feedback** on query effectiveness
- 5. Provide support for collaboration and team working

Russell-Rose & MacFarlane, EARS 2020: The 3rd International Workshop on ExplainAble Recommendation and Search2020

Sharing, reuse, best practice

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





Summary

- Reimagining 'advanced search'
 - Eliminate syntactic error, make semantics transparent
 - Provide support for reuse, optimisation, best practice
- Key transformations:
 - Procedural → declarative
 - Static → interactive
 - Monolithic → executable
 - Strings \rightarrow objects
- Try it for yourself: https://www.2dsearch.com
 - Tutorial videos: https://www.2dsearch.com/help

Thank you

Dr. Tony Russell-Rose, FBCS CITP CEng

Founder: 2Dsearch, Director: UXLabs

Senior Lecturer, Goldsmiths

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