Institute of Education

UCL

# Machine learning and automation in reviews is now a reality; but do we yet know how to use these technologies?

**James Thomas, Sergio Graziosi, Jeff Brunton**
**Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)**
**Social science research unit, Department of Social Science**
**UCL Institute of Education**
**University College London**

ePPI CENTRE

# Declaration of interests and funding

- James Thomas is co-lead of the Cochrane ‘Transform’ project, which is implementing some of the technologies discussed here. He also directs development & management of EPPI-Reviewer, the EPPI-Centre’s software for systematic reviews.

- Parts of this work funded by: Cochrane, JISC, Medical Research Council (UK), National Health & Medical Research Council (Australia), Wellcome Trust, Bill & Melinda Gates Foundation. All views expressed are my own, and not necessarily those of these funders.

# Objectives

- To introduce technologies for automating parts of the review process including: study selection; risk of bias assessment; and synthesis;
- For participants to try these technologies for themselves; and
- To discuss methodological issues concerning their use.

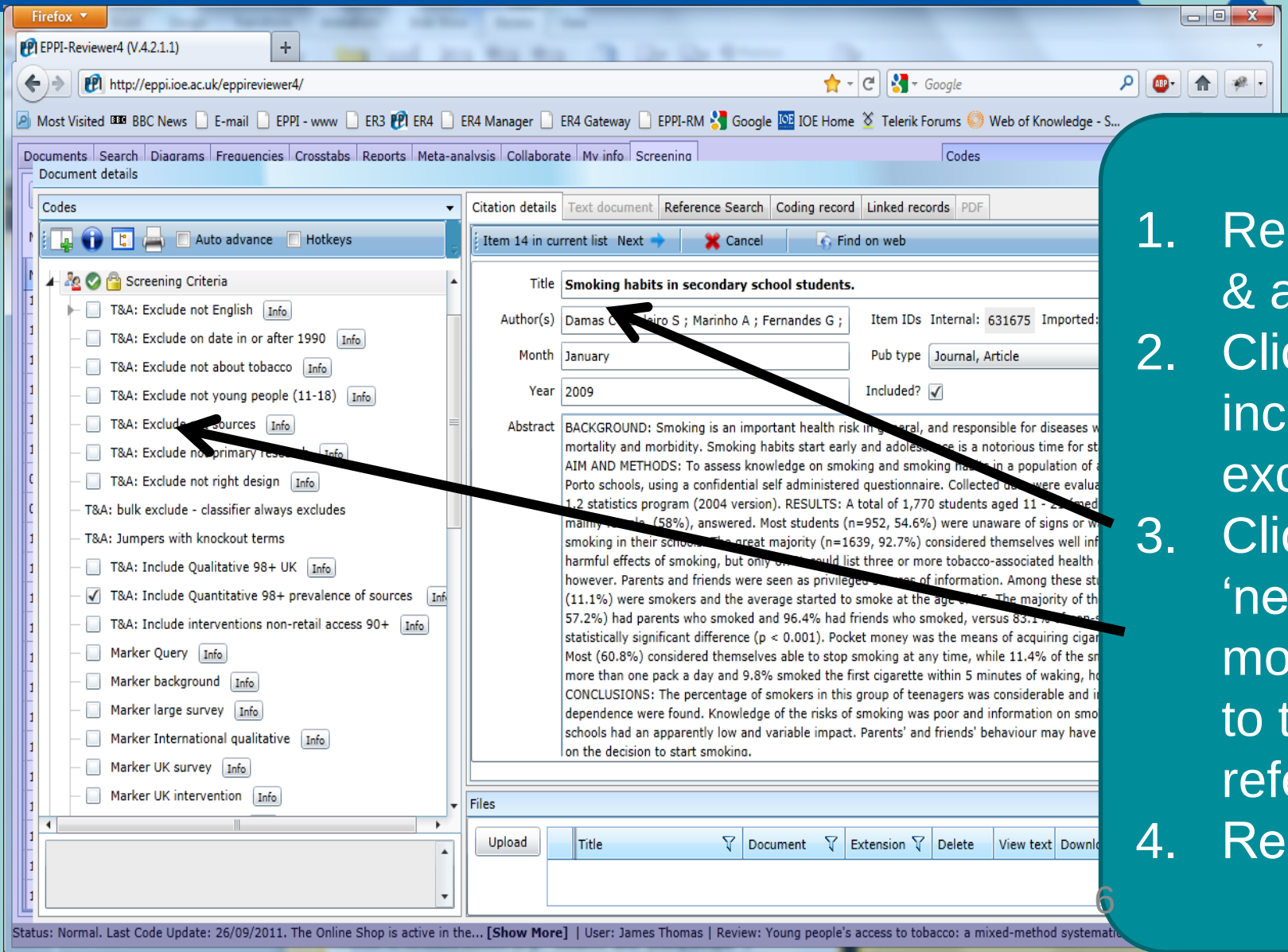
Structure:

- Presentation (many slides…) + questions / discussion
- Experimentation with online tools and discussion (small groups)
- Whole group feedback (if there's time / need)
- (**http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3677**)

# Automation in systematic reviews – what can be done?

- Study identification:
  - Citation screening
  - Updating reviews
  - RCT classifier
- Mapping research activity
- Search strategy development
- Data extraction
  - Risk of Bias assessment
  - Other study characteristics
  - Extraction of statistical data
- Synthesis and conclusions

# Reducing workload during citation screening




1. Read title & abstract
2. Click include / exclude
3. Click ‘next’ and move on to the next reference
4. Repeat…

Institute of Education

ePPI CENTRE

# Citation screening

- Has received most R&D attention
- Diverse evidence base; difficult to compare evaluations
- ‘semi-automated’ approaches are the most common
- Possible reductions in workload in excess of 30% (and up to 97%)

O'Mara-Eves *et al. Systematic Reviews* 2015, **4**:5
http://www.systematicreviewsjournal.com/content/4/1/5

SYSTEMATIC REVIEWS

RESEARCH Open Access

## Using text mining for study identification in systematic reviews: a systematic review of current approaches

Alison O'Mara-Eves[1], James Thomas[1*], John McNaught[2], Makoto Miwa[3] and Sophia Ananiadou[2]

**Abstract**

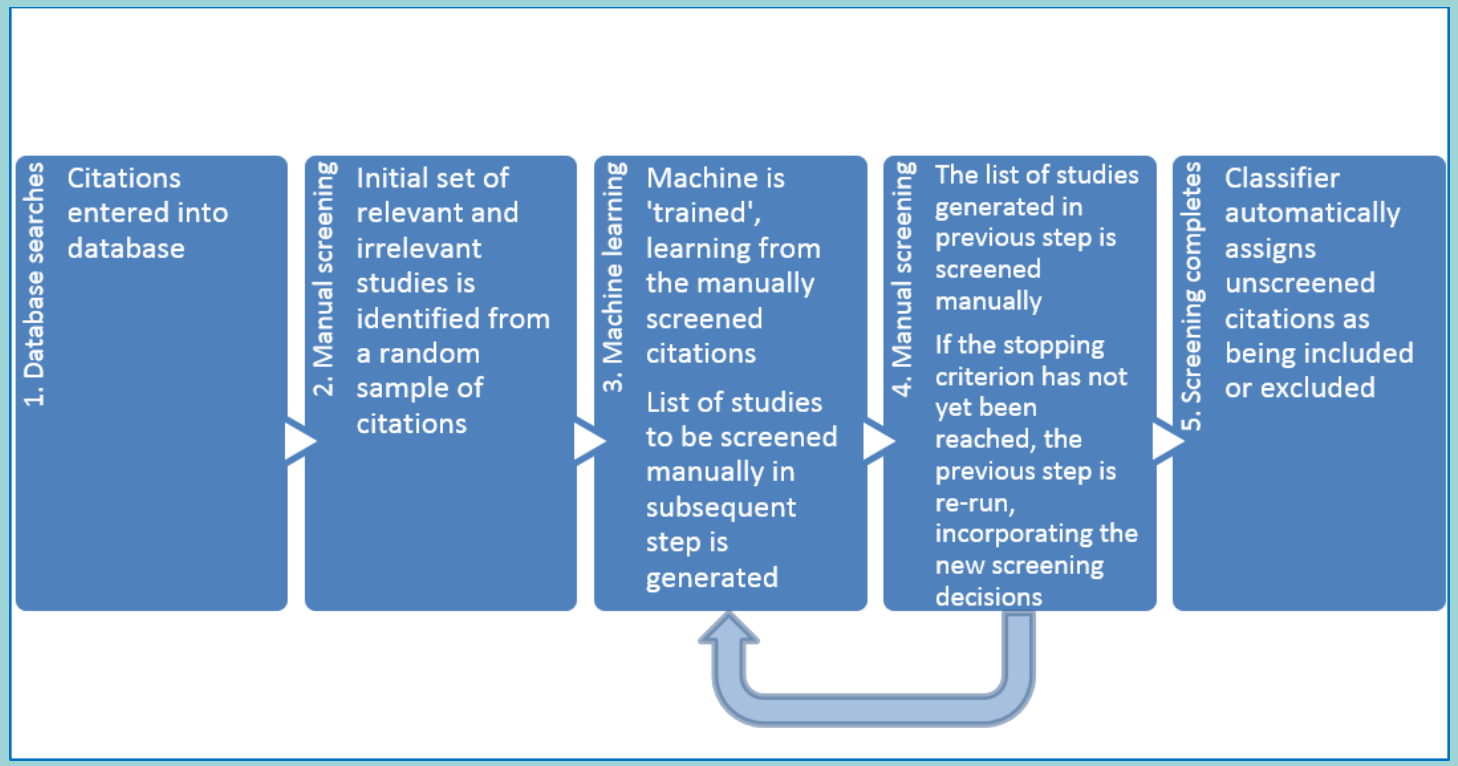
**Background:** The large and growing number of published studies, and their increasing rate of publication, makes the task of identifying relevant studies in an unbiased way for inclusion in systematic reviews both complex and time consuming. Text mining has been offered as a potential solution: through automating some of the screening process, reviewer time can be saved. The evidence base around the use of text mining for screening has not yet been pulled together systematically; this systematic review fills that research gap. Focusing mainly on non-technical issues, the review aims to increase awareness of the potential of these technologies and promote further collaborative research between the computer science and systematic review communities.

**Methods:** Five research questions led our review: what is the state of the evidence base; how has workload reduction been evaluated; what are the purposes of semi-automation and how effective are they; how have key contextual problems of applying text mining to the systematic review field been addressed; and what challenges to

## Summary of conclusions

- Screening prioritisation
  - ‘safe to use’
- Machine as a ‘second screener’
  - Use with care
- Automatic study exclusion
  - Highly promising in many areas, but performance varies significantly depending on the domain of literature being screened

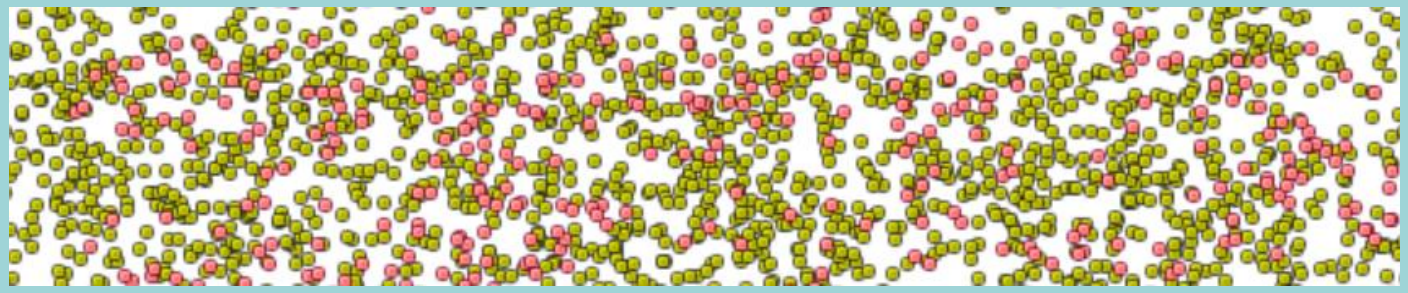
# How the machine learns…




And it can work quite well…

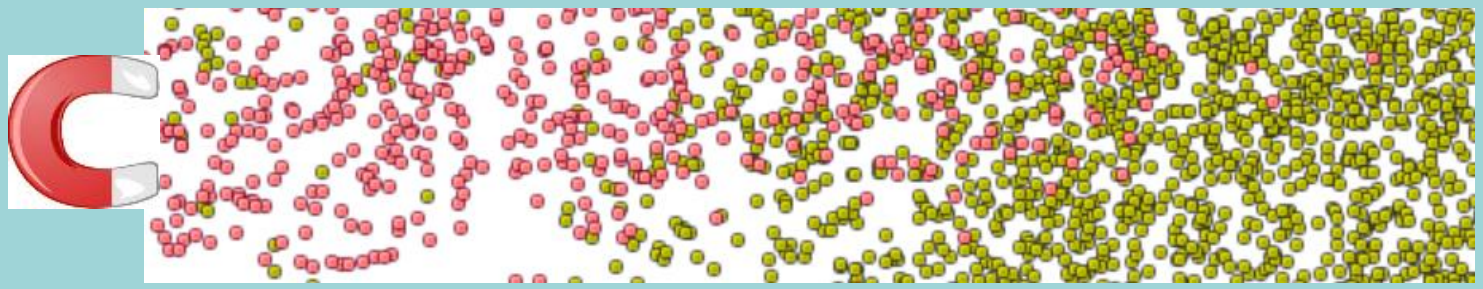
# Screening prioritisation: Changing the distribution of studies

Traditional screening

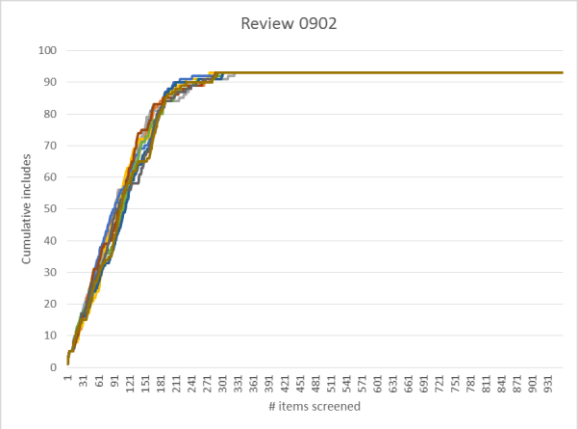


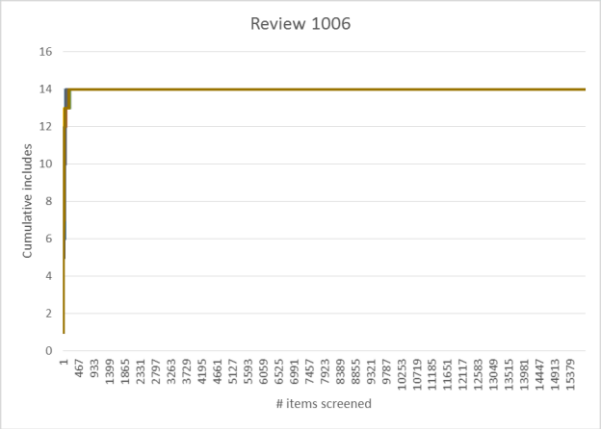
Screening process (red = eligible study)

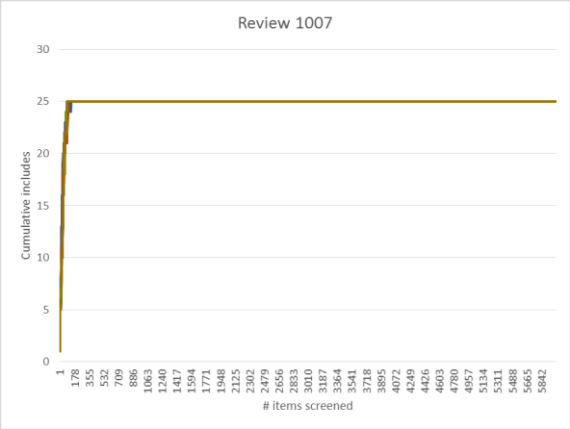
Screening aided by text mining

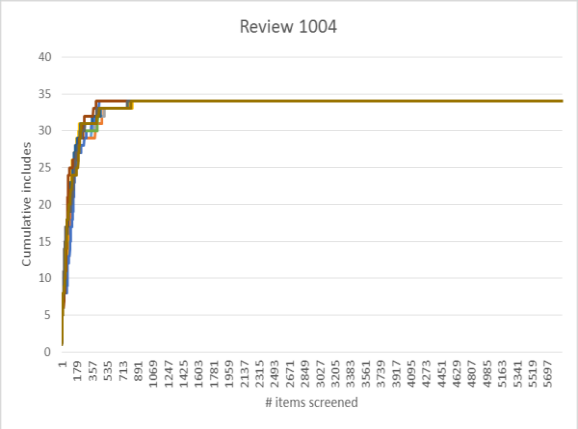


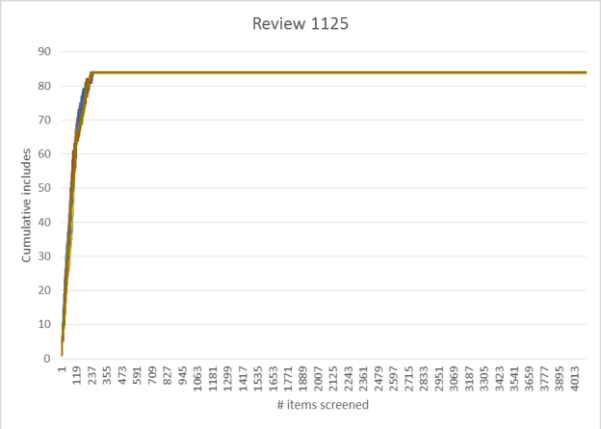
# e.g. reviews from Cochrane Heart Group

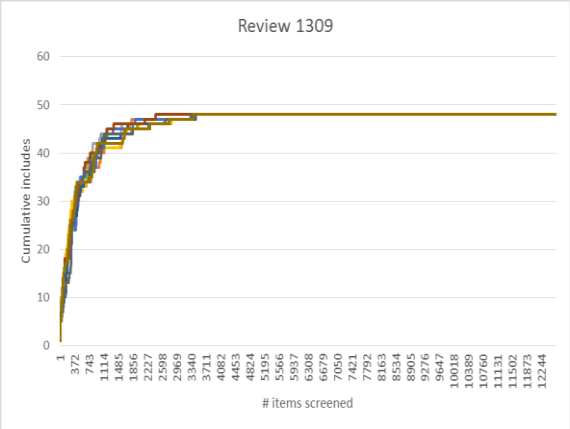


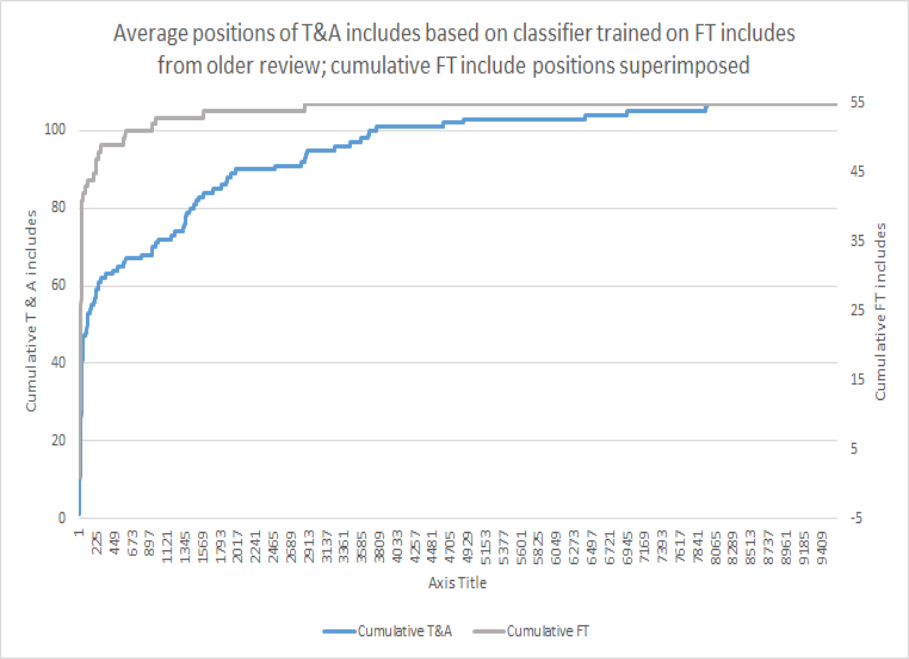



# Updating existing reviews

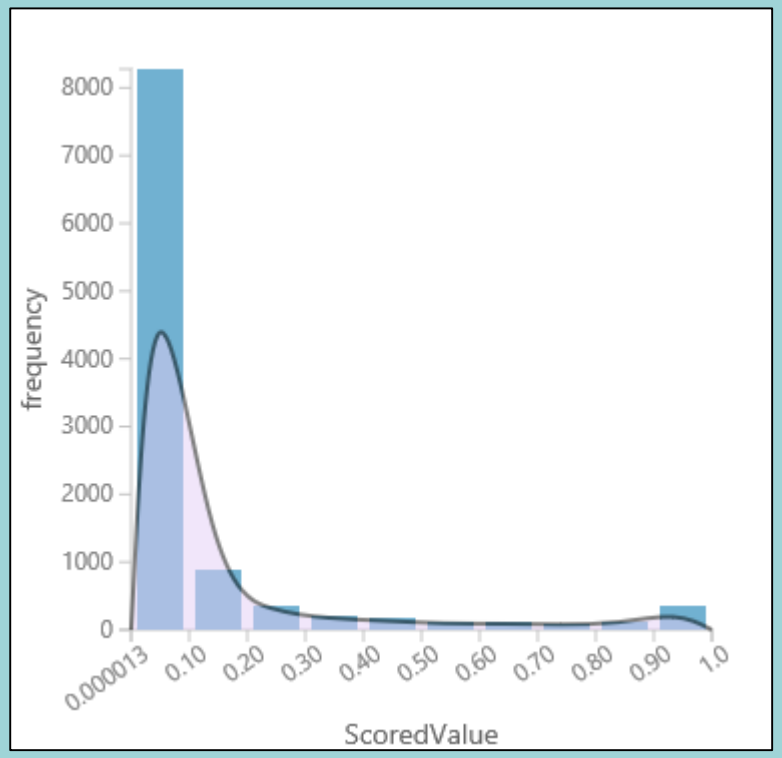



Weightman A, Thomas J, Baker P, Lovie-Toon Y, Francis D, O'Mara-Eves A (2014) Text mining for screening efficiency? Testing within a Cochrane public health review. Poster presented at Cochrane Colloquium 2014, Hyderabad

# RCT Classifier

# RCT Classifier

- ‘Trained’ on more than 280,000 human classifications from Cochrane Crowd
- Recall = 99.879% if all citations > 0.1 are screened manually; 60% workload saving??



# Institute of Education

http://localhost/WcfHostPortal/EppiReviewer4.a

EPPI-Reviewer4 (V.4.6.1.2)

Documents | Search | Diagrams | Frequencies | Crosstabs | Reports | Meta-analysis | Collaborate | My info | Screening

New search | Refresh search list | Delete selected | Combine | AND | OR | NOT (included) | NOT (excluded)

| | Title | Created by | Date | Hits | Select |
|---|---|---|---|---|---|
| 176 | Items classified according to model: RCT | James Thomas | 10/18/2016 | 22 | Select |
| 175 | Items classified acc | | | | |
| 174 | Coded with: EXC1- | | | | |
| 173 | Coded with: EXC1- | | | | |
| 172 | "low income" AND | | | | |
| 171 | low income team | | | | |
| 170 | food club | | | | |
| 166 | moynihan | | | | |
| 165 | roe | | | | |
| 164 | kennedy | | | | |
| 163 | friends with food | | | | |
| 162 | hodgson | | | | |
| 161 | McGlone | | | | |
| 160 | white and carlin | | | | |
| 159 | white | | | | |
| 158 | carter | | | | |
| 157 | lloyd | | | | |
| 156 | Dutton | | | | |
| 155 | cook for Life | | | | |
| 154 | cook4Life | | | | |
| 153 | gregg | | | | |
| 152 | Barton | Rebecca Rees | 4/21/2011 | 15 | Select |
| 151 | Symon | Rebecca Rees | 4/21/2011 | 5 | Select |
| 150 | 149 AND 57 | James Thomas | 4/19/2011 | 1400 | Select |
| 149 | Coded with: To screen pages 1 and 3 | James Thomas | 4/19/2011 | 1400 | Select |

Machine learning classifier

The machine learning classifier built into EPPI-Reviewer enables you automatically to assign a code to new items based on those you have assigned to others. There are two stages to the process: first the classifier 'learns' to apply the code by building a model (which is saved for future use); and then the model is applied to new items. One 'pre-built' classifier is available: the 'RCT model', which will automatically identify RCTs for you; it has been trained on more than 280,000 records screened by the Cochrane Crowd.
(This function enables you to build a linear classifier from a bag of words representation of your studies, using the scikit-learn python library.)

**Stage 1: build the model**

Learn to apply this code

Distinguish from this code

Name for your model

Build model

(Go straight to stage 2 if you are applying the RCT model.)

**Stage 2: apply the model** | Refresh model list

| Title | Applies | Compared wi | Precision | Recall |
|---|---|---|---|---|
| Check screening | Include | Exclude | 0.17 | 1.00 |

Apply above selected model

Apply RCT model

Apply to all items in review

Apply to items with this code

Apply to items from this source

Apply model

Codes

TestingTM

Screening criteria

al

ing allocation

G clusters

my Tool

code set

ct tool

Codes | Sources | Review statistics

Status: Normal. The Online Shop is now active in the Account Manager pages. It... **[Show More]** | User: James Thomas | Review: Cooking Skills Review (Expired)

# Institute of Education

UCL

http://localhost/WcfHostPortal/EppiReviewer4.a

EPPI-Reviewer4 (V.4.6.1.2)

Documents | Search | Diagrams | Frequencies | Crosstabs | Reports | Meta-analysis | Collaborate | My info | Screening

29 documents loaded.

Showing: Items classified according to model: RCT

Filter:

| | | | Author | Title | Year | Score |
|---|---|---|---|---|---|---|
| Go | | I | Horner S | Enhancing Asthma Self-Management in Rural School-Aged Children: A Randomized Controlled Trial | 2015 | 99 |
| Go | | I | McGhan S | A children's asthma education program: Roaring Adventures of Puff (RAP), improves quality of life | 2010 | 99 |
| Go | | I | Bruzzese | Feasibility and impact of a school-based intervention for families of urban adolescents with asthma: results from a randomized pil | 2008 | 98 |
| Go | | I | Horner SI | Improvement of rural children's asthma self-management by lay health educators | 2008 | 98 |
| Go | | I | Cicutto L | Breaking the access barrier: evaluating an asthma center's efforts to provide education to children with asthma in schools | 2005 | 97 |
| Go | | I | Bruzzese | Reducing Morbidity And Urgent Health Care Utilization In Urban Pre-adolescents With Asthma: Results Of A Randomized Control T | 2010 | 96 |
| Go | | I | Kintner El | Randomized clinical trial of a school-based academic and counseling program for older school-age students | 2009 | 96 |
| Go | | I | McWhirte | Can schools promote the health of children with asthma? | 2008 | 96 |
| Go | | I | Patterson | A cluster randomised intervention trial of asthma clubs to improve quality of life in primary school children: the School Care and A | 2005 | 96 |
| Go | | I | Shah S ; | Effect of peer led programme for asthma education in adolescents: cluster randomised controlled trial | 2001 | 95 |
| Go | | I | Clark NM | An evaluation of asthma interventions for preteen students | 2010 | 93 |
| Go | | I | MeGhan S | Evaluation of an education program for elementary school children with asthma | 2003 | 86 |
| Go | | I | Bruzzese | Effects of a school-based intervention for urban adolescents with asthma. A controlled trial | 2011 | 84 |
| Go | | I | Gerald LB | Outcomes for a comprehensive school-based asthma management program | 2006 | 81 |
| Go | | I | Clark NM | Effects of a comprehensive school-based asthma program on symptoms, parent management, grades, and absenteeism | 2004 | 78 |
| Go | | I | Butz A ; F | Rural children with asthma: Impact of a parent and child asthma education program | 2005 | 77 |
| Go | | I | Joseph CL | A web-based, tailored asthma management program for urban African-American high school students | 2007 | 76 |
| Go | | I | Cicutto L | A randomized controlled trial of a public health nurse-delivered asthma program to elementary schools | 2013 | 75 |
| Go | | I | Guglani L | Exploring the impact of elevated depressive symptoms on the ability of a tailored asthma intervention to improve medication adh | 2013 | 74 |
| Go | | I | Levy M ; | The efficacy of asthma case management in an urban school district in reducing school absences and hospitalizations for asthma | 2006 | 74 |
| Go | | I | Mandhane | A child's asthma quality of life rating does not significantly influence management of their asthma | 2010 | 74 |
| Go | | I | Clark NM | A trial of asthma self-management in Beijing schools | 2005 | 70 |
| Go | | I | Joseph CL | Feasibility Of Web-based Asthma Education For Urban Teenagers: Intervention Compliance And Computer Access [Abstract] | 2010 | 70 |
| Go | | I | Henry RL | Randomized controlled trial of a teacher-led asthma education program | 2004 | 63 |
| Go | | I | McCann D | A controlled trial of a school-based intervention to improve asthma management | 2006 | 60 |
| Go | | I | Kintner El | Effectiveness of a school-based academic asthma health education and counseling program on fostering acceptance of asthma in | 2014 | 57 |

1 Page 1 of 1

Codes

- Source
- James's codes
- Screening Criteria
- Admin codes
- EPPI-Centre Health pron
- Mapping Coding Tool
- Mapping tool allocations
- Surveys n=100
- Jeff 3.1 check
- YP tobacco sources - Da

Codes | Sources | Review statistics

Status: Normal. The Online Shop is now active in the Account Manager pages. It... **[Show More]** | User: James Thomas | Review: Young people's access to tobacco: a mixed-method systematic review.

# Developing search strategies

# Text recognition/text analytics

- Text analysis (frequency counts, phrases or nearby terms in text, or statistical frequency across corpus)
- Term extraction and automatic clustering (ranked list of words or phrases from a combination of linguistic and statistical analyses)

## Applications:

- Provides a rapid overview of words/phrases or controlled terms in sample, depending on tools used.
- Identifying words or phrases that not considered
- Identify unwanted search terms

# Termine: Automatic Term Identification

http://www.nactem.ac.uk/software/termine/

- Termine automatically identifies and ranks terms according to their importance to the citation list

The software tools and services which NaCTeM supplies allow researchers to apply text mining techniques to problems within their specific areas of interest - examples of these tools are highlighted below. In addition to providing services , the Centre is also involved in , and makes significant contributions to , the text mining research community both nationally and internationally in initiatives such as Europe PubMed Central.

| Rank | Term | Score |
|---|---|---|
| 1 | text mining | 3 |
| 2 | text mining research community | 2 |
| 3 | datum mining system | 1.584962 |
| 3 | natural language processing | 1.584962 |

Frantzi, Katerina, Sophia Ananiadou, and Hideki Mima. "Automatic recognition of multi-word terms:. the C-value/NC-value method." International Journal on Digital Libraries 3.2 (2000): 115-130.

# Text recognition/text analytics

Sample N=52

Using TD*IDF

| Term | Score |
|---|---|
| ID | 76.61 |
| people | 44.67 |
| resident | 39.81 |
| adult | 38.26 |
| woman | 37.08 |
| older people | 35.73 |
| menopause | 32.58 |
| participant | 29.33 |
| client | 28.07 |
| life | 27.35 |
| older adult | 26.21 |
| death | 25.76 |
| older person | 25.76 |
| age | 23.62 |
| active ageing | 19.76 |
| retirement | 19.55 |
| person | 18.72 |
| population | 18.14 |
| end-of-life care | 17.95 |
| future planning | 17.12 |
| ageing adult | 15.8 |

Using Termine

| | |
|---|---|
| intellectual disability | 148.896545 |
| aged care | 14.3125 |
| end-of-life care | 10 |
| intellectual disability nurse | 6.33985 |
| community-based aged care | 6.33985 |
| active ageing | 6 |
| moderate id | 6 |
| down syndrome | 5 |
| future planning | 5 |
| residential aged care | 4.754888 |
| residential aged care facility | 4 |
| learning disability | 4 |
| moderate intellectual disability | 3.169925 |
| future plan | 3 |
| developmental disability | 3 |
| id group home staff | 2 |
| mild id | 2 |
| future care | 2 |
| death method | 2 |
| hospice care | 2 |
| elderly people | 2 |
| profound id | 2 |
| late life | 2 |
| age-related illness | 2 |
| retirement option | 2 |
| community-based aged care results | 2 |
| nursing home | 2 |
| normal ageing | 2 |

# Identifying search terms for:
# Research on care and support needs for ageing populations of adults with learning disabilities

52 items relevant to topic

↓

Analysed with both Termine and TD*IDF

↓

Selected population terms related to older people, ageing, aged care.
Re-ran search to find 44/52 items.

↓

Scanned remaining 8 records manually,

↓

Revisited term lists to identify additional terms:

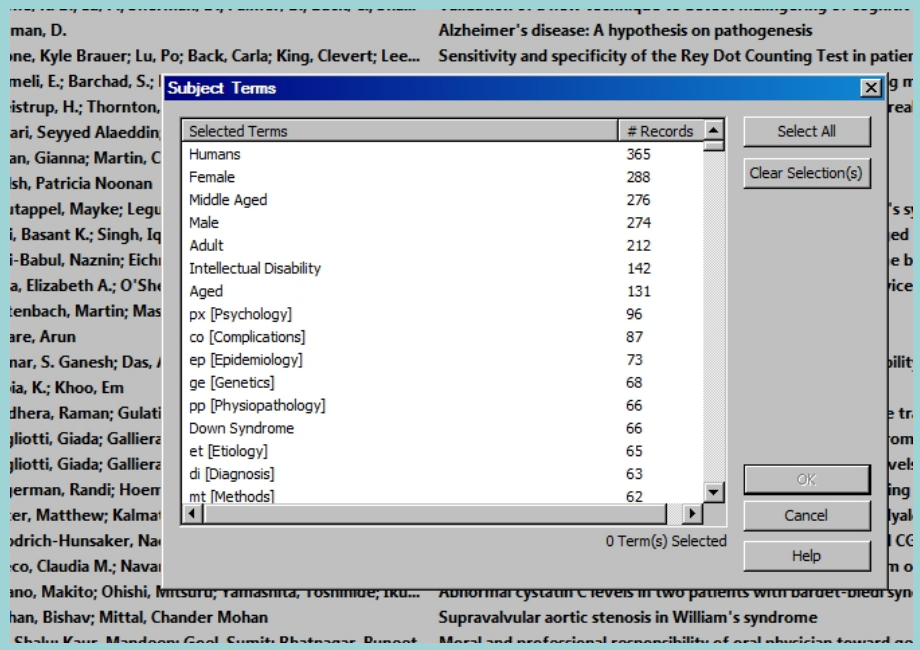
`Future planning, planning for the future, Aged related illness, menopausal`

↓

Incorporated into existing search strategy

# Rapid overview of controlled terms

Analysing controlled terms in Endnote:




N=500 items to analyse search subset

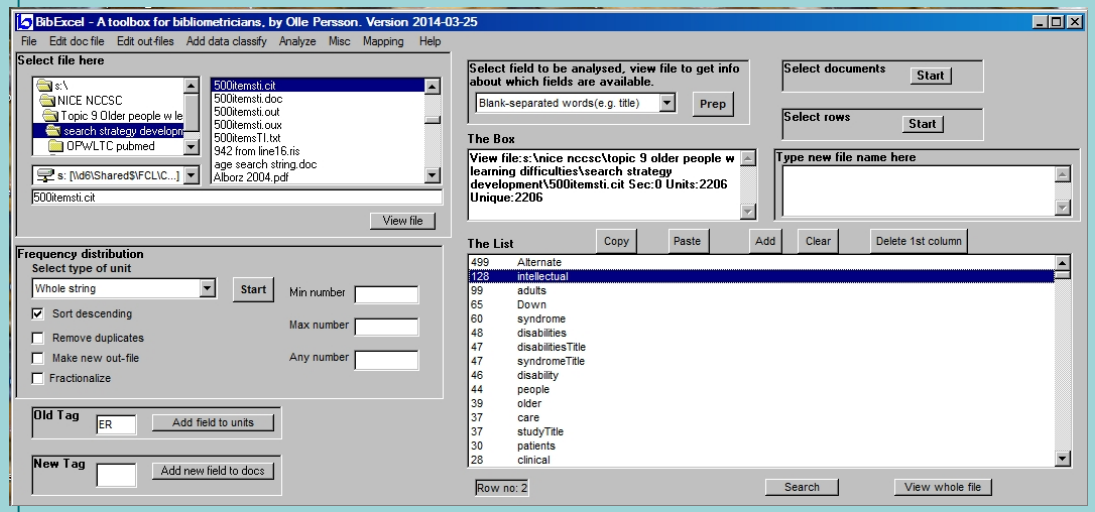
Method in: Hayman, S and Shaheem, Y (2014). Smart Searching: Logical Steps to Building and Testing Your Literature Search. CareSearch Palliative Care Knowledge Network.http://sites.google.com/site/smartsearchinglogical/home

# Rapid overview of words per citation

Word frequency per citation:

In Bibexcel analysed titles/abstracts from 1,000 items in test search - count per citation, found:

Gene (96), Premutation (59), protein (57), FMR1 (57), mutation (51), FXTAS (44)

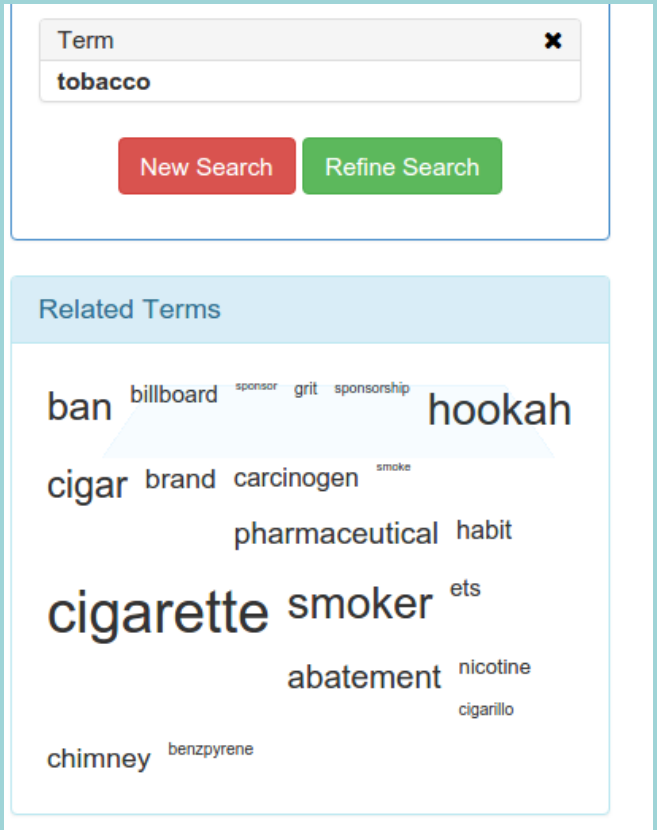



Directory of tools:

http://www.tapor.ca/

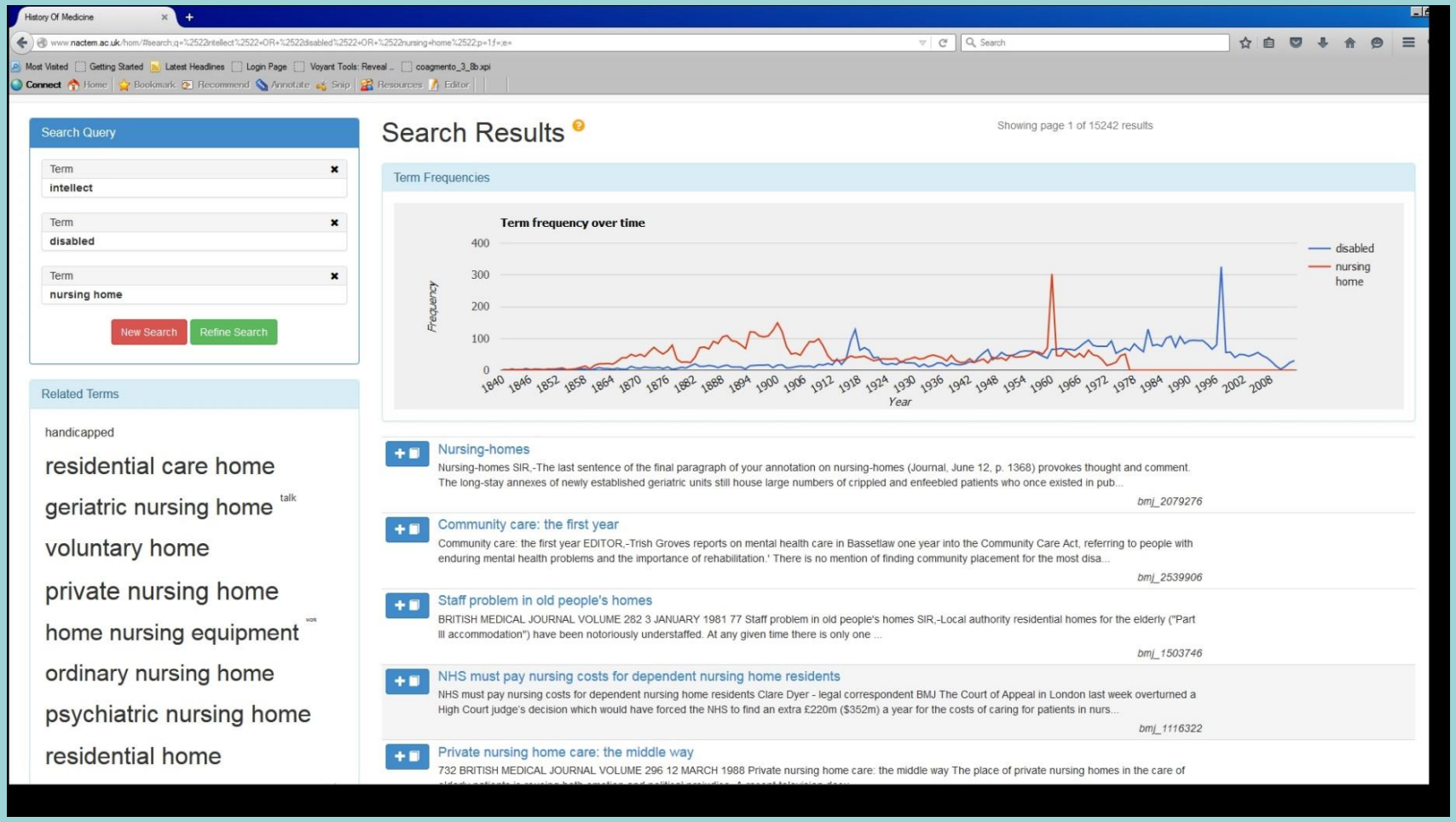
# Query expansion
## http://nactem.ac.uk/hom

- Augment query with synonyms, related terms, orthographic variations etc.



# Query expansion

http://nactem.ac.uk/hom



# Mapping research activity

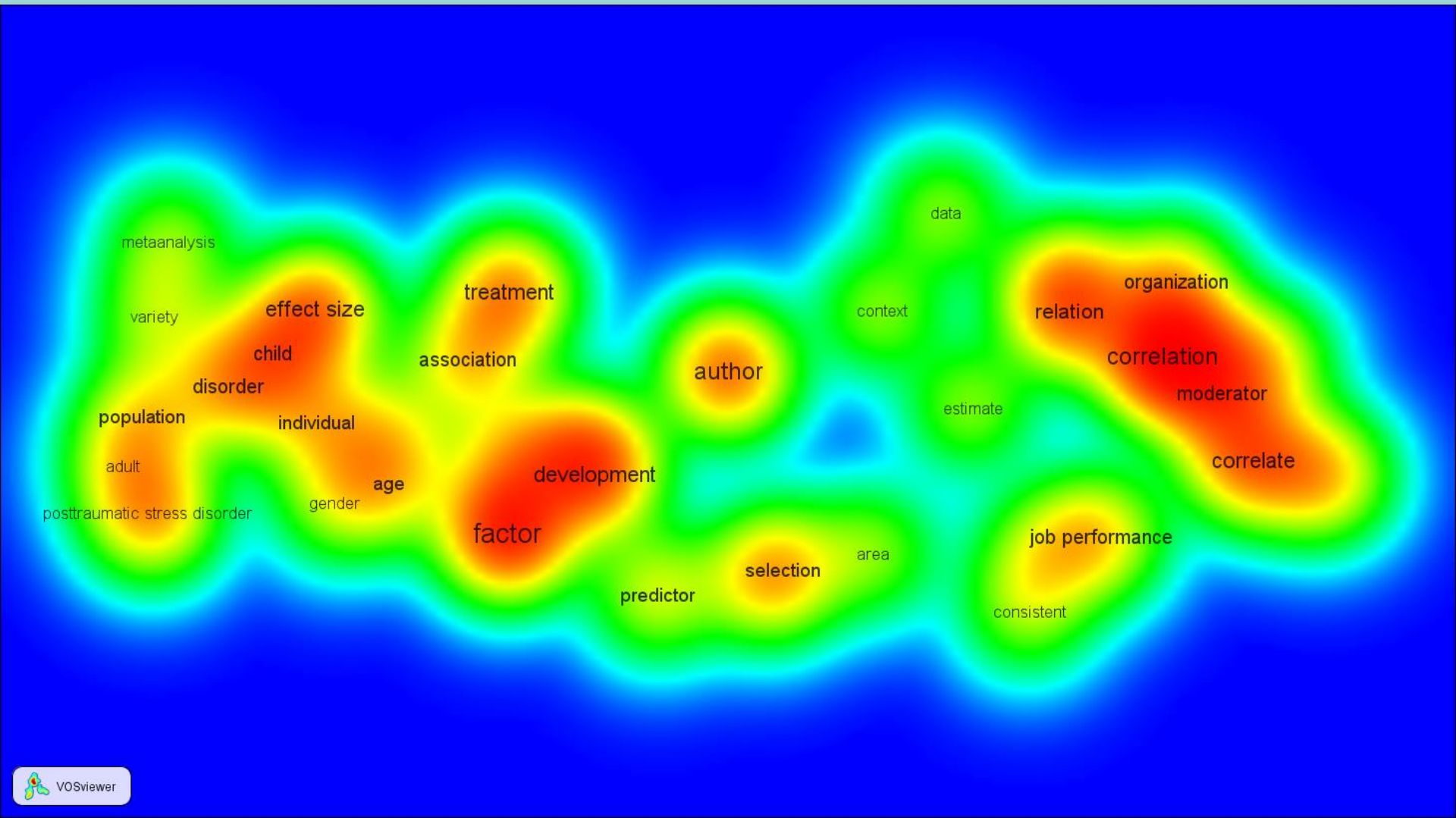
# Technologies for identifying sub-sets of citations

- Different families of techniques
  - Fairly simple approaches which examine term frequencies to group similar citations
  - More complex approaches, such as Latent Dirichlet Allocation (LDA)
- Identifies groups of documents which use similar combinations of terms
- The difficult part is finding good labels to describe the clusters
  - But are labels always needed?
- Visualisations are often incorporated into tools

Institute of Education

ePPI CENTRE

UCL

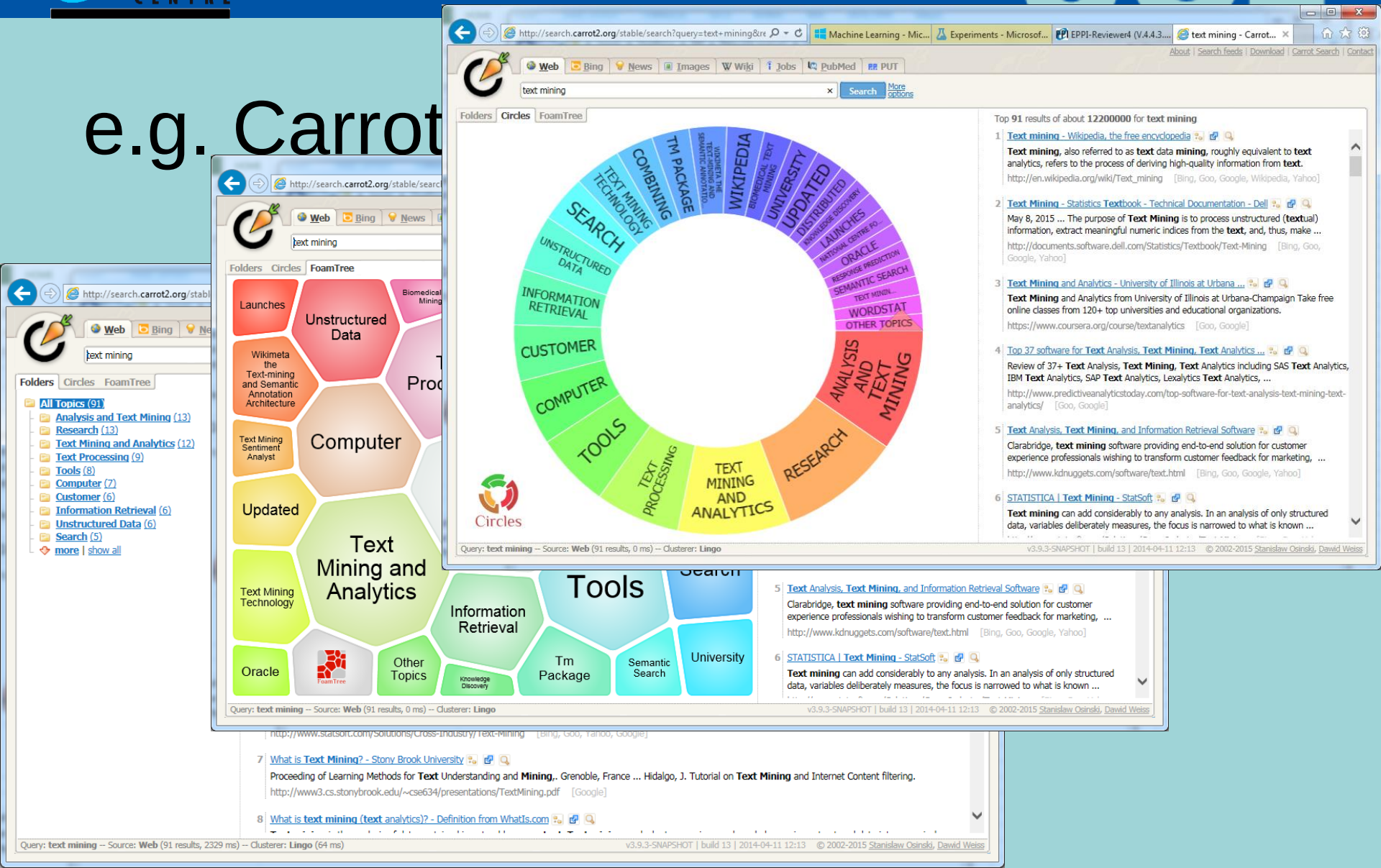


Institute of Education

ePPI CENTRE

UCL

# e.g. Carrot



# “Topic modelling”

CRD_lda_topic50_AB - Notepad

File Edit Format View Help

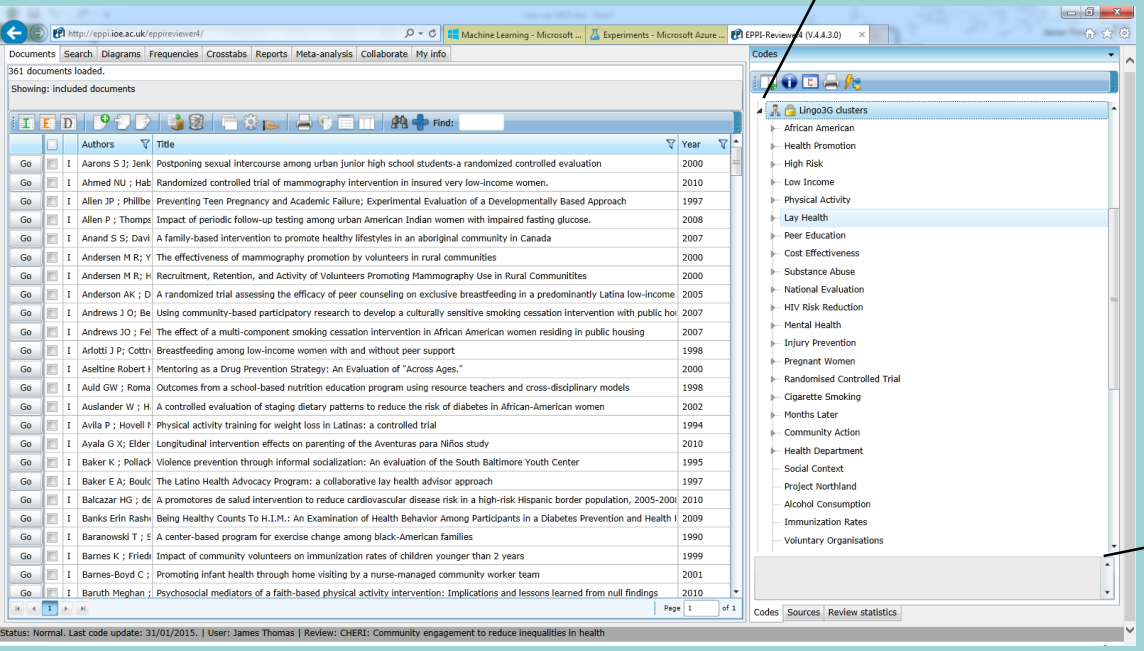
```
0.148*use + 0.050*drug + 0.038*substance + 0.036*alcohol + 0.022*drugs + 0.020*users +
0.016*marijuana + 0.016*among + 0.015*cocaine + 0.014*cannabis0.052*sexual + 0.038*hiv +
0.031*sex + 0.029*risk + 0.020*use + 0.017*drug + 0.015*partners + 0.015*among + 0.015*men +
0.014*infection0.069*intake + 0.050*dietary + 0.034*food + 0.025*diet + 0.018*energy +
0.015*vitamin + 0.014*consumption + 0.014*fat + 0.014*fruit + 0.013*vegetables0.101*men +
0.071*women + 0.026*body + 0.019*mass + 0.018*fat + 0.017*age + 0.015*associated +
0.012*index + 0.012*association + 0.009*associations0.013*compliance + 0.011*volume +
0.008*immigrants + 0.007*salivary + 0.007*epilepsy + 0.006*volumes + 0.006*immigrant +
0.006*doctors + 0.006*conventional + 0.006*acamprosate0.094*blood + 0.078*pressure +
0.054*hypertension + 0.025*systolic + 0.020*bp + 0.015*diastolic + 0.013*diabetic + 0.012*mm
+ 0.010*hypertensive + 0.010*mmhg0.110*risk + 0.072*disease + 0.057*factors + 0.040*heart +
0.039*cardiovascular + 0.033*coronary + 0.029*stroke + 0.022*chd + 0.019*smoking +
0.014*hypertension0.019*markers + 0.016*stress + 0.015*cortisol + 0.014*tissue + 0.014*muscle
+ 0.013*inflammation + 0.012*growth + 0.012*inflammatory + 0.011*oxidative + 0.010*visceral
0.019*media + 0.016*internet + 0.012*information + 0.011*ra + 0.011*messages + 0.010*online +
0.009*content + 0.009*mdma + 0.008*advertising + 0.008*incontinence0.084*suicide +
0.042*suicidal + 0.030*athletes + 0.030*attempts + 0.026*sports + 0.021*ideation +
0.017*sport + 0.017*attempt + 0.014*attempted + 0.010*team0.142*patients + 0.013*treatment +
0.011*clinical + 0.011*patient + 0.010*hospital + 0.009*therapy + 0.007*cases + 0.007*pain +
0.006*medication + 0.006*diagnosis0.076*95 + 0.065*ci + 0.038*risk + 0.025*odds +
0.021*confidence + 0.020*associated + 0.018*ratio + 0.018*interval + 0.015*association +
0.013*adjusted0.072*gambling + 0.023*problem + 0.021*gamblers + 0.020*pathological + 0.010*pd
```

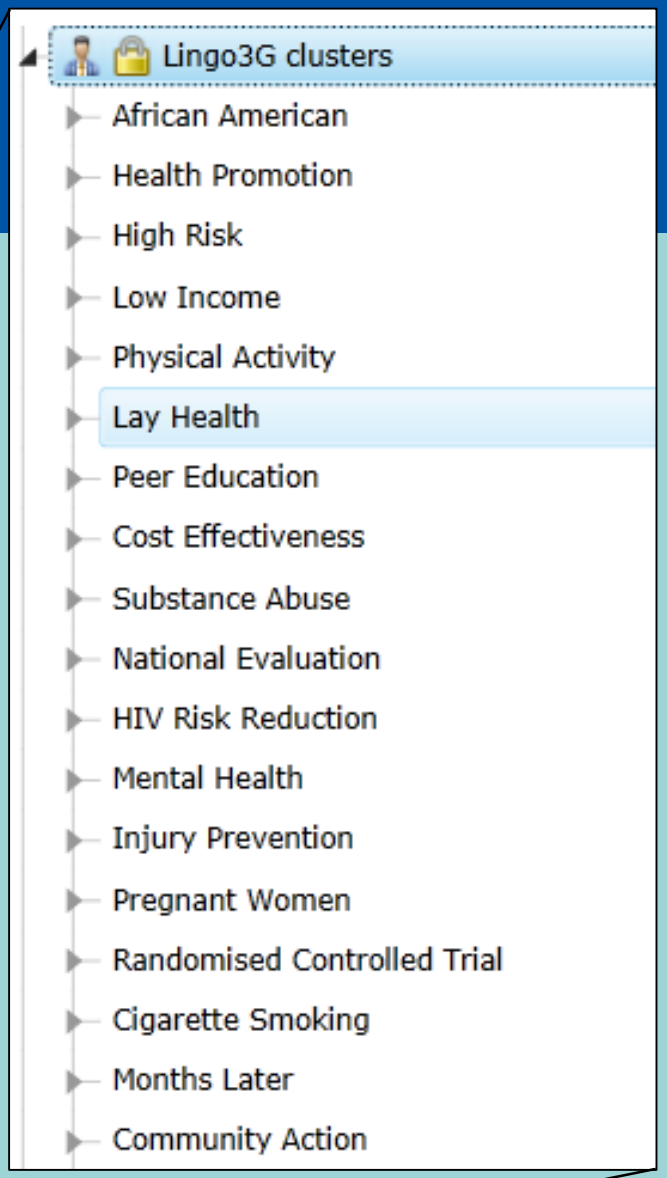
| | A | B | C | D | E | F | G | H | I | J | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Included? | Topic0 | Topic1 | Topic2 | Topic3 | Topic4 | Topic5 | Topic6 | Topic7 | Topic8 | Topi |
| 2 | No | 7.395112 | 5.946818 | 5.040004 | 4.744282 | 4.253583 | 3.83581 | 2.417316 | 3.087711 | 4.860644 | 3.0 |
| 3 | Yes | 14.07189 | 2.137577 | 4.443276 | 5.034399 | 0.7431 | 2.299016 | 1.26992 | 1.574956 | 7.010434 | 4.6 |

Institute of Education

ePPI CENTRE

# Used in a review about community engagement to improve health amongst disadvantaged groups



- Quick map codes
  - Health topic area
  - Progress+
  - Type of community engagement
  - Outcomes
    - Blood pressure
    - Body mass index
    - Cost effectiveness
    - Fruit & veg intake
    - Exclusive breastfeeding
    - Immunization rates
    - Low birth weight
    - Mortality
    - Safe sex behaviours
    - Smoking cessation
  - Age group
    - Older adults
    - Children/ young people
    - Workers/ employees

- Quick map codes
  - Health topic area
  - Progress+
  - Type of community engagement
    - Capacity building
    - Community action/ support
    - Community involvement
    - Community partnership
    - Community mobilization
    - Community organisations
    - Computer services
    - Cultural competence
    - Health advocacy
    - Lay health
    - Neighbourhood management
    - Neighbourhood Wardens Scheme
    - Outreach programme
    - Participatory research
    - Peer counseling
    - Peer education
    - Peer leaders
    - Peer support
    - Social networks
    - Task Force
  - Outcomes

- Quick map codes
  - Health topic area
    - Antenatal care
    - Asthma prevention / treatment / management
    - Breastfeeding
    - Cardiovascular disease
    - Cancer prevention
    - Cancer screening / detection
    - Care of older people
    - Child abuse prevention
    - Child illness and ill health
    - Diabetes prevention/ management
    - Disabilities & chronic illness
    - Healthy eating
    - HIV Prevention / risk reduction
    - Housing
    - Hypertension
    - Immunisation
    - Injury prevention
    - Mental health
    - Neighbourhood renewal
    - Obesity prevention / weight reduction
    - Partner violence
    - Physical activity
    - Public health/ Health promotion/ prevention

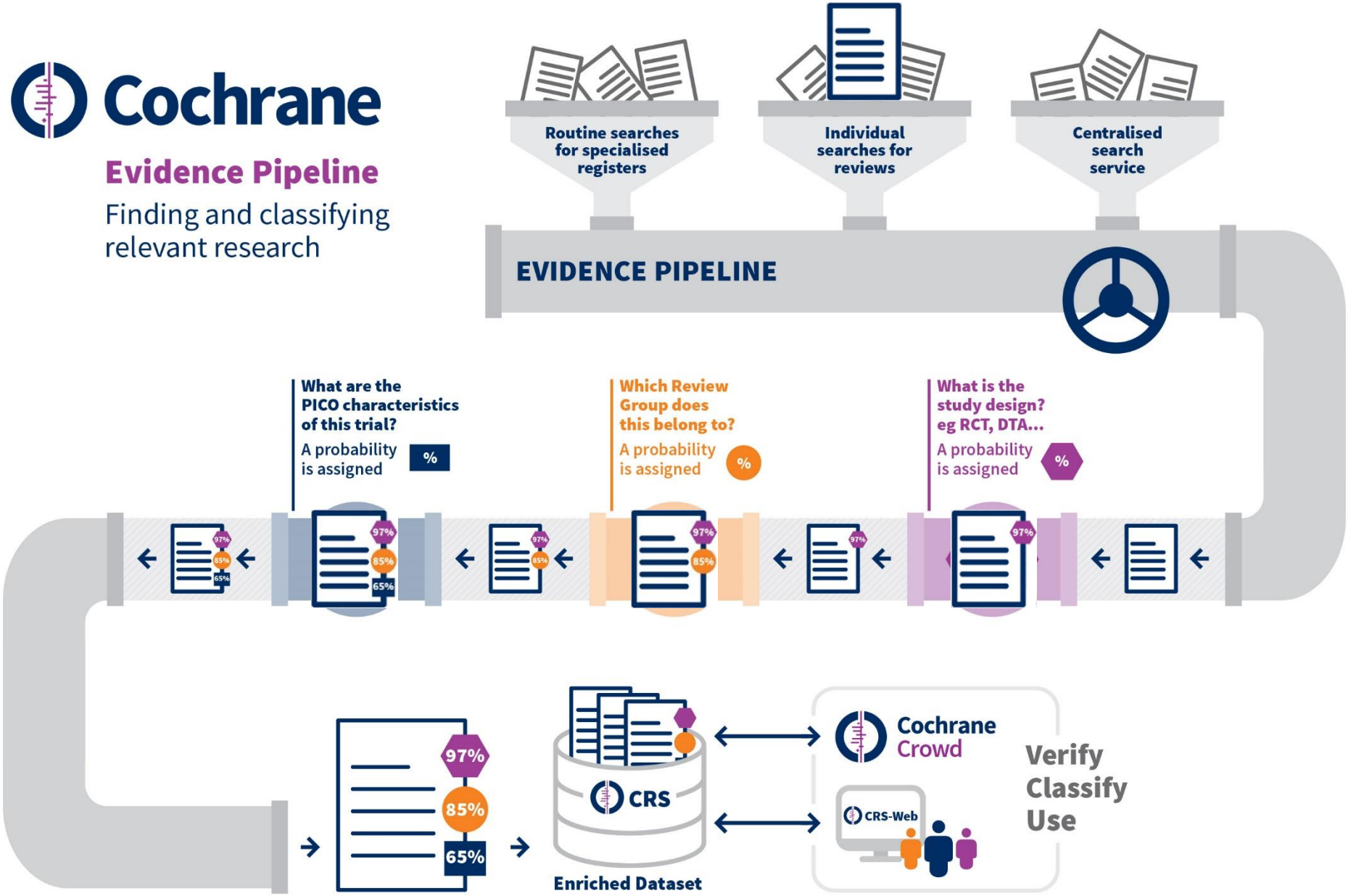
With a little organisation and user interaction, we constructed a workable ‘map’

# Creating databases of research prospectively

# Cochrane

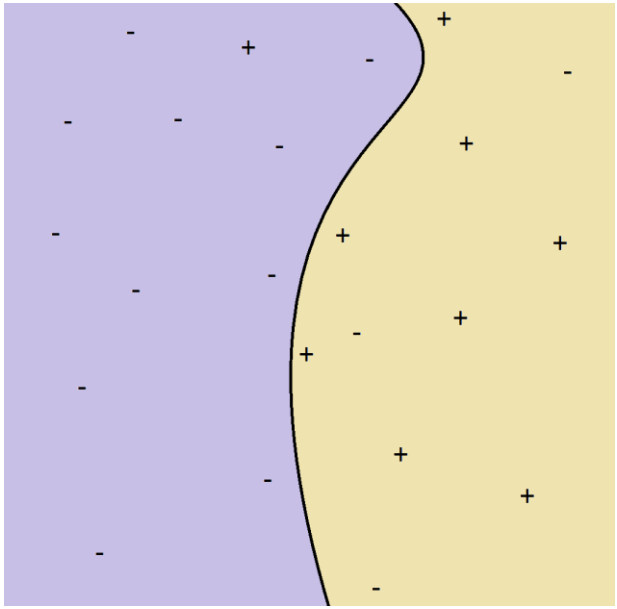
## Evidence Pipeline

Finding and classifying relevant research



Project Transform

# How does it work?



Building machine classifiers

Cochrane

Project Transform

# 1. A dictionary and index are created

- First, the key terms in the studies are listed (ignoring very common words
- Second, the studies are indexed against the list of terms
  - (the resulting matrix can be quite large)
- Next…

**e.g. We have two studies – one is an RCT, and one isn't an RCT**

Study 1 **Effectiveness** of **asthma self-care interventions**: a **systematic review** (not an RCT)

Study 2 **Effectiveness** of a **self-monitoring asthma intervention**: an **RCT** (an RCT)

| **RCT?** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| **1** | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |

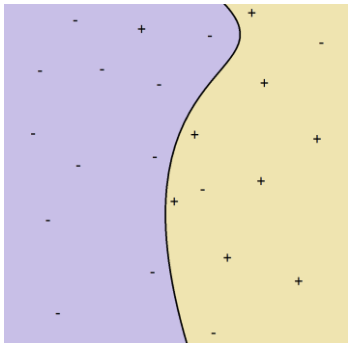
Cochrane

Project Transform

# 2. A statistical model is built

The matrix is used to create a statistical model which is able to distinguish between the two classes of document (e.g. between RCTs and non-RCTs)

| RCT? | Effectiveness | asthma | self | care | interventions | systematic | review | monitoring | intervention | RCT |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |



Cochrane

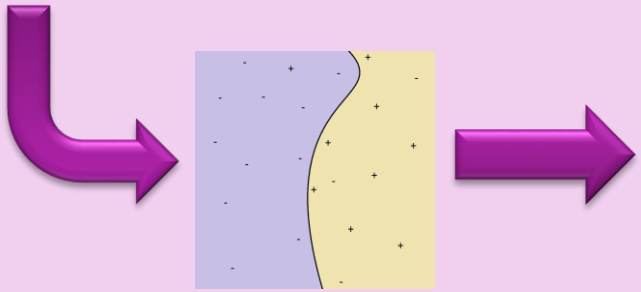
Project Transform

# 3. The model is applied to new documents

- New citations are indexed against the previously generated list of terms
- The resulting matrix is fed into the previously generated model
- And the model will assign a probability that the new document is, or is not a member of the class in question

e.g. The **effectiveness** of a school-based **asthma** management programme: an RCT

| Effectiveness | asthma | self | care | interventions | systematic | review | monitoring | intervention | RCT |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |



**93%**

Cochrane

Project Transform

# The new CRS-Web

CRS Web (online)

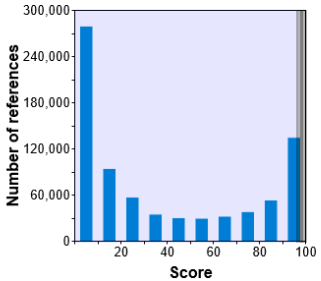
crsdemo.metaxis.com/index.php#Search

Search | Simple | MeSH | Classifier | Saved | Tracking

## Classifier search

Records that have been through the classifier have probabilities assigned to them to indicate how likely they are to have certain properties, like whether they are of interest to a review group, or whether they are likely to be an RCT. Choose the classifier model you are interested in, set the model parameters and click Search to find the records

RCT




Approximately 32129 records that are between 99 and 100 percent likely to be of interest

Search

You can find your records that are currently being processed by the classifier by seaching for INPROCESS:CLASSIFIER

Find those records

Cochrane Register of Studies

Anna Noel-Storr [DEMENTIA] | logout

Dashboard | Records | Import | Journals | CT.GOV | Reports | To do

Search | Layout 1 | Layout 2 | Layout 3 | Layout 4 | Deduplication | New reference | New study | Users | Help

Search results

Export | Find and replace | Add to folder | Add to... | Remove from... | Add to marked list

(399 records) Page 1 of 8

| # | Title | Author |
|---|---|---|
| 1 | Cognitive effects of treating obstructive sleep apnea in Alzheimer's disease: a randomized controlled study | Ancoli-Israel S // Palmer BW // Cooke |
| 2 | Efficacy of galantamine in probable vascular dementia and Alzheimer's disease combined with cerebrovascular disease: a rando… | Erkinjuntti T // Kurz A // Gauthier S // |
| 3 | Donepezil improved memory in multiple sclerosis in a randomized clinical trial | Krupp LB // Christodoulou C // Melvil |
| 4 | A randomized, 26-week, double-blind, placebo-controlled trial to evaluate the safety and efficacy of galantamine in the treatme… | Auchus A |
| 5 | A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Donepezil Study Group | Rogers SL // Farlow MR // Doody RS / |
| 6 | A Controlled, Double-Blind, Randomized Pilot Clinical Trial of Hydroxysafflor Yellow a on Cognitive Function in Patients With Vas… | Tian J |

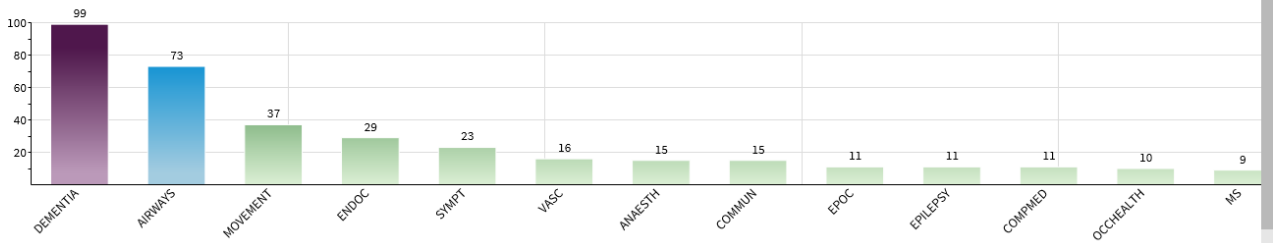
Record

Fields | Duplicates | Links | Reviews | Classifier | Files | Audit | CENTRAL | REGISTER

The bar chart below shows the classifier scores for this record. Scores are presented in the range 0 -1 00 where higher scores mean a higher likelihood that the record is of interest to the group. You can tell a group about this record if it doesn't already have it in it's segment by clicking the bar for that group.

In register | In segment | Not in segment | Not relevant to my group

There is a **99%** likelihood that this record is an RCT [Confirm this is **not** an RCT] [Confirm this **is** an RCT]



Cochrane

Project Transform

Home | How it works | About Epistemonikos | Languages | 7 | My Epistemonikos | dperezrada@gmail.com

# Collection: Direct oral anticoagulants for thromboprophylaxis after orthopedic surgery

Owner: dperezrada@gmail.com

Update recommendations | Delete Collection

Related (26) | Included (119) | Excluded (59) | Machine Recommendations (24)

Primary Study

**Single intravenous administration of TB-402 for the prophylaxis of venous thromboembolism after total knee replacement: a dose-escalating, randomized, controlled trial.**

AUTHORS » Verhamme P , Tangelder M , Verhaeghe R , Ageno W , Glazer S , Prins M , Jacquemin M , Büller H , TB-402 Study Group

JOURNAL » Journal of thrombosis and haemostasis : JTH

YEAR » 2011

LINKS » Pubmed , DOI

This article is included in 1 Systematic Review

Include | Exclude

Systematic Review

**Rivaroxaban for thromboprophylaxis after orthopaedic surgery: pooled analysis of two studies.**

AUTHORS » Fisher WD , Eriksson BI , Bauer KA , Borris L , Dahl OE , Gent M , Haas S , Homering M , Huisman MV , Kakkar AK , Kälebo P , Kwong LM , Misselwitz F , Turpie AG

JOURNAL » Thrombosis and haemostasis

YEAR » 2007

LINKS » Pubmed

Without references

Include | Exclude

Unclassified

**Fondaparinux for prevention of venous thromboembolism in major orthopedic surgery.**

AUTHORS » Tran AH , Lee G

JOURNAL » The Annals of pharmacotherapy

YEAR » 2003

LINKS » Pubmed , DOI

Include | Exclude

Primary Study

First | Previous | 1 | 2 | 3 | Next | Last

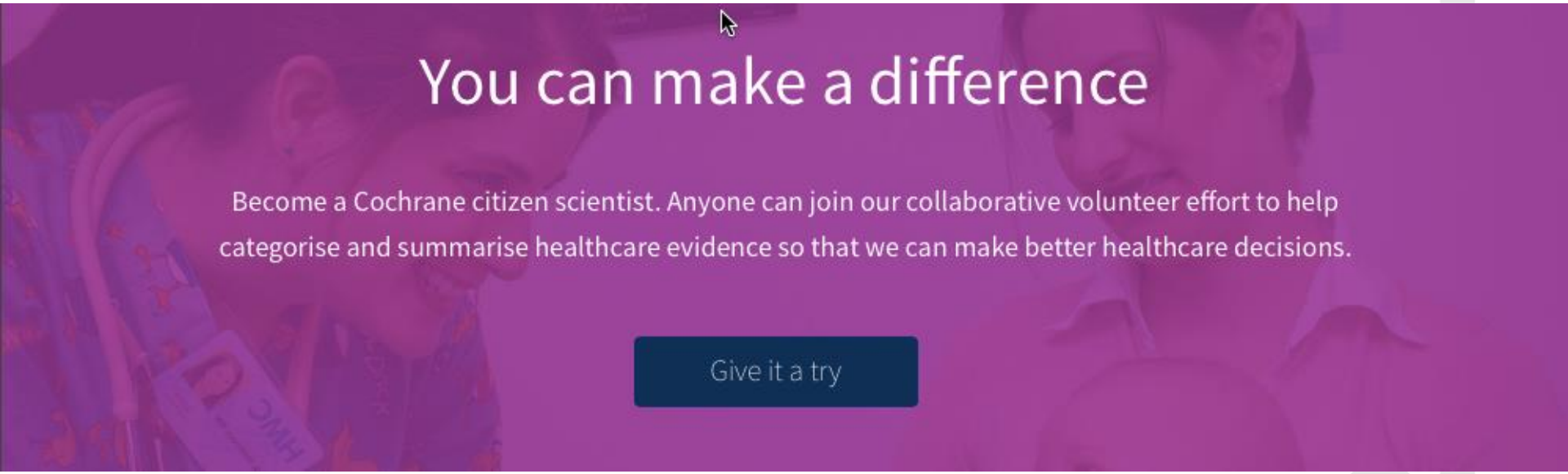
Abstract | About this article

Despite current guidelines recommendations about anticoagulant prophylaxis, many studies have shown an high venous thromboembolism (VTE) incidence in patients undergoing total hip and knee arthroplasty. A number of anticoagulants are currently available, but they have some limitations that affect their applicability and consequently their effectiveness. Several new oral anticoagulants (NOACs) have been developed in an attempt to overcome these limitations. Apixaban is a NOAC that selectively inhibits the coagulation factor Xa; it is approved for the prevention of VTE after total hip replacement and total knee replacement surgery. This review examines the results of main trials designed to test efficacy and safety of apixaban in major elective orthopedic surgery.

Cochrane

Project Transform

# Today




Cochrane Crowd
crowd.cochrane.org

# Risk of bias assessment

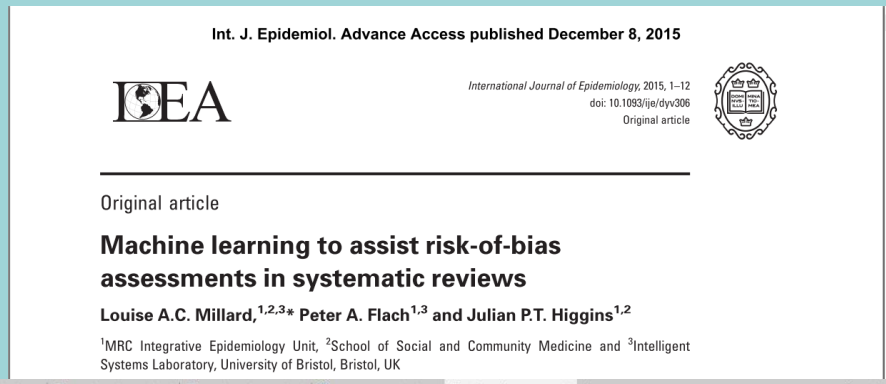
Institute of Education

ePPI CENTRE

UCL

# Risk of Bias assessment

- Emerging area; e.g.
  - RobotReviewer
  - Millard, Flach and Higgins
- Tools can accomplish two purposes:
  - 1. identify relevant text in the document
  - 2. automatically assess risk of bias
- Can perform very well though authors do not yet suggest well enough to replace humans



Int. J. Epidemiol. Advance Access published December 8, 2015

IJEA

*International Journal of Epidemiology*, 2015, 1–12
doi: 10.1093/ije/dyv306
Original article

Original article

**Machine learning to assist risk-of-bias assessments in systematic reviews**

**Louise A.C. Millard,[1,2,3]* Peter A. Flach[1,3] and Julian P.T. Higgins[1,2]**

[1]MRC Integrative Epidemiology Unit, [2]School of Social and Community Medicine and [3]Intelligent Systems Laboratory, University of Bristol, Bristol, UK

# Data extraction

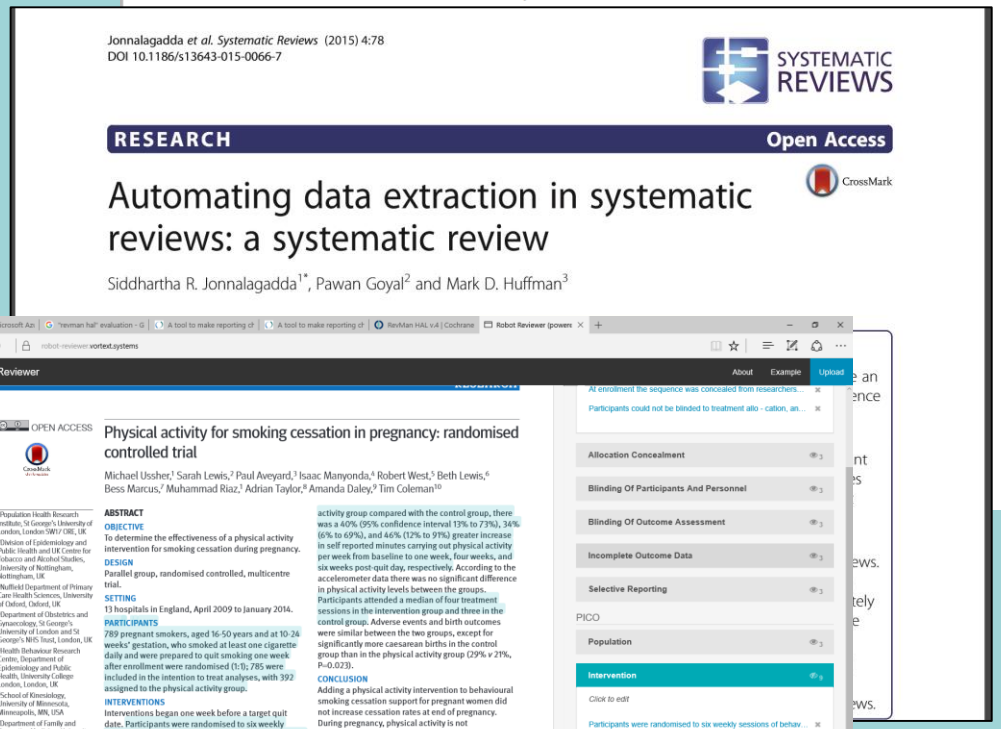
Institute of Education

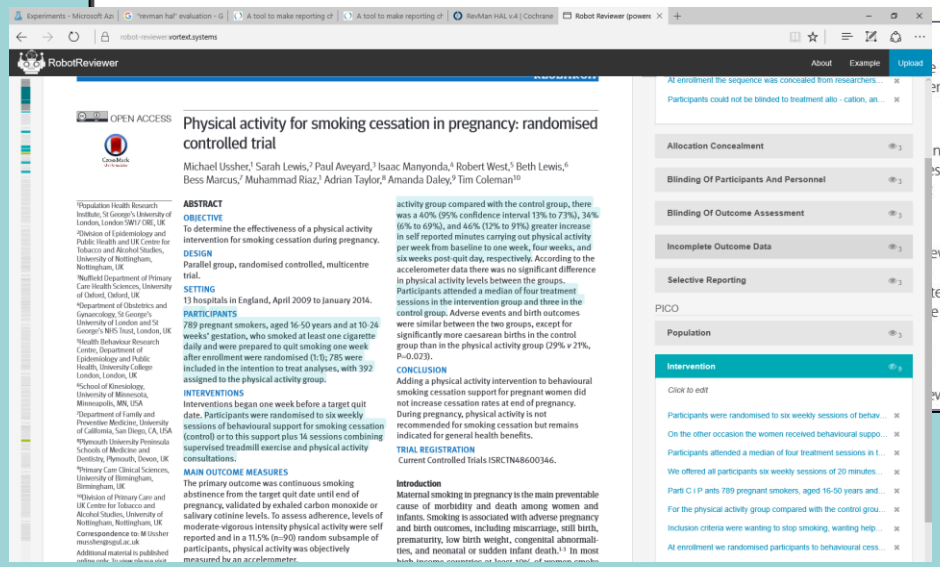
ePPI CENTRE

UCL

# Data extraction

- RobotReviewer can identify phrases relating to study PICO characteristics
- ExaCT extracts trial characteristics (e.g. eligibility criteria)
- Systematic review found that no unified framework yet exists
- More evaluative work is needed on larger datasets
- Further challenges include extraction of data from tables and graphs

# Synthesis


enterprise
infrastructure
cards
text mining
solution
stakeholder

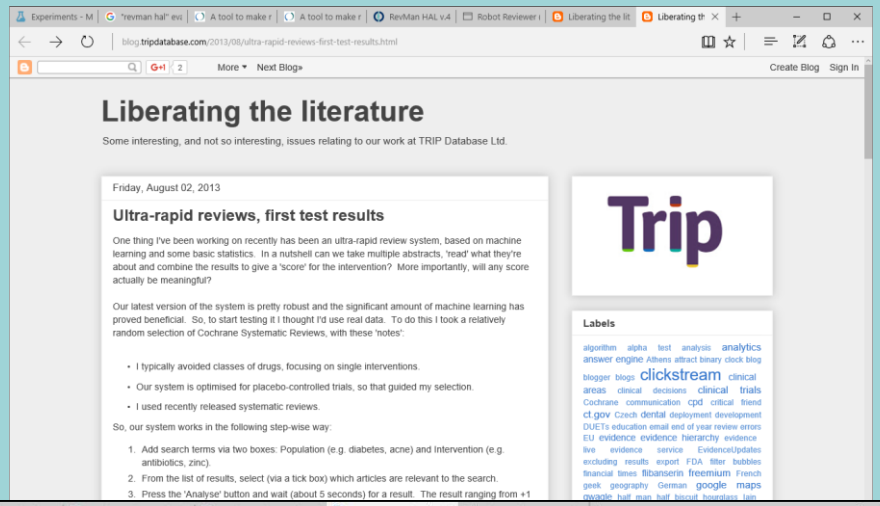
Institute of Education

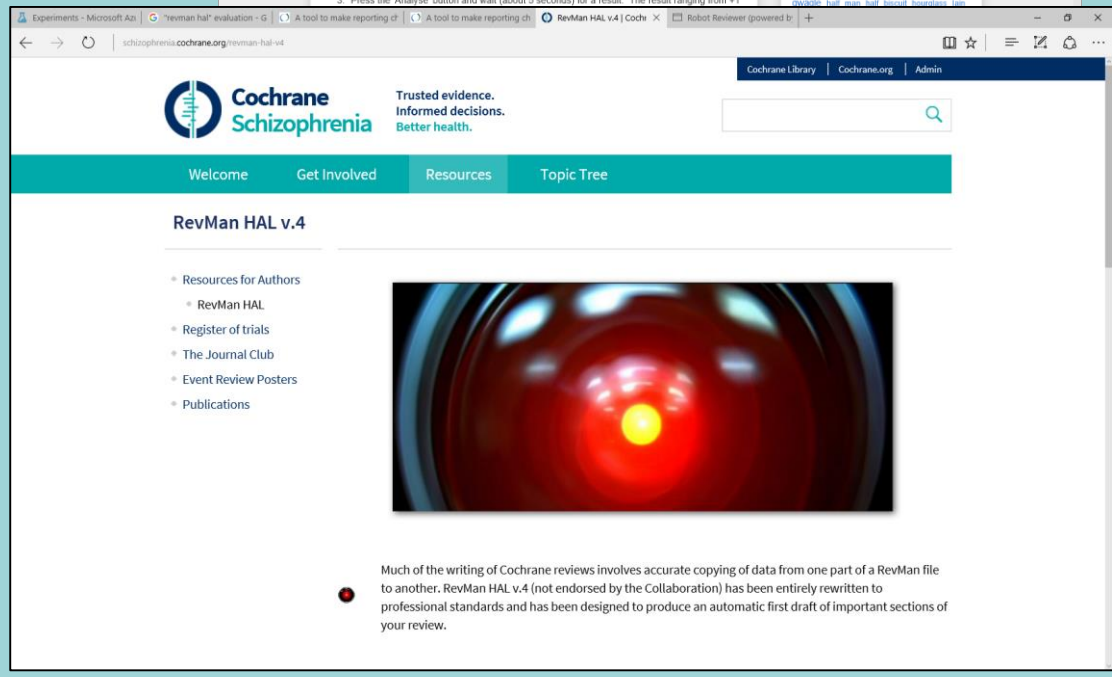
ePPI CENTRE

UCL

# Synthesis and conclusions

- Summarisation and synthesis of text is an active area for development in computer science
- Many hurdles to overcome before this technology can be used routinely
- Some systems automate parts of the process



# Tools to try

This page: http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3677

Classification: RCT Classifier; and ‘custom’ classifier

- EPPI-Reviewer: http://eppi.ioe.ac.uk.

Workload reduction during citation screening (using ‘active learing’)

- Rayyan systematic reviews tool: http://rayyan.qcri.org/reviews/5
- Microsoft Azure Machine Learning (James can demo for those interested)

Identifying sub-sets of citations (clustering)

- Carrot2 search: http://search.carrot2.org/stable/search
- Topic modelling: http://nbviewer.jupyter.org/github/bmabey/pyLDAvis/blob/master/notebooks/GraphLab.ipynb#topic=12&lambda=0.84&term=

Search strategy development

- Termine: http://www.nactem.ac.uk/software/termine/
- NaCTeM History of Medicine: http://nactem.ac.uk/hom

# Discussion: in small groups

**Discuss methodological issues that the use of these technologies raise**

- which tools do you already use?
- what challenges do you face using existing tools?
  Is recall of 99.879% for a workload reduction of 60% reasonable?
- what challenges do you think the new technologies would raise?
- what do you think would help you to adopt the new technologies?
- are there situations when using these tools might threaten the reliability of the review?

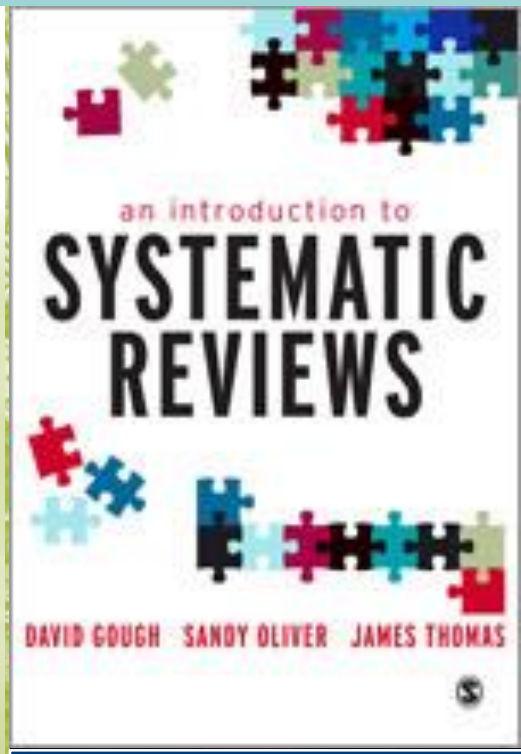
Institute of Education

# Thank you

SSRU website: http://www.ioe.ac.uk/ssru/
SSRU's EPPI website: http://eppi.ioe.ac.uk

Email
j.thomas@ioe.ac.uk
c.stansfield@ioe.ac.uk
a.omara-eves@ioe.ac.uk

**The EPPI-Centre is part of the Social Science Research Unit at the UCL Institute of Education, University College London**




**EPPI-Centre**
Social Science Research Unit
Institute of Education
University of London
18 Woburn Square
London WC1H 0NR

Tel +44 (0)20 7612 6397
Fax +44 (0)20 7612 6400
Email eppi@ioe.ac.uk
Web eppi.ioe.ac.uk/