Current Strategies for Updating Systematic Reviews

A webcast hosted by AIR’s Center on KTDRR
January 26, 2022

James Thomas
Under construction
Outline

• A new search paradigm for updating systematic reviews and maps?
• Case study: the COVID-19 map of research maintained by EPPI-Centre, University College London and CRD, University of York
A new search paradigm?

• Two enablers:
  • Increasingly open bibliographic data
  • New automation technologies and tools
Finding eligible studies

All research reports

Stored in multiple databases / sites

Boolean searches retrieve most likely relevant into local database

Records deduplicated and manually screened for eligibility
Why is it like this?

- Various organisations have indexed research in their area.
- It costs money to index research.
  - The business model (thus far) has been to charge for access to bibliographic databases.
  - There’s been financial benefit to keeping a monopoly hold over the content of databases (i.e. you can’t find a given set of records by searching elsewhere).
- This has led to the picture we see of numerous, subject-specific databases jealously guarding their content (and charging £££ for access).
Movement towards open and more comprehensive sources

- Open Alex / CrossRef / Dimensions… (Google Scholar to some extent) are publishing bibliographic data at scale.
- Microsoft Academic – now discontinued – played a major role.
- Bibliographic data are becoming commoditized (you can get the same content from multiple sources).
- You can ingest (essentially) ALL the world’s published research into a local database.

An open and comprehensive catalog of scholarly papers, authors, institutions, and more.

Inspired by the ancient Library of Alexandria, OpenAlex is an index of hundreds of millions of interconnected entities across the global research system. We're 100% free and open source, and offer access via a web interface, API, and database snapshot.
Commoditized bibliographic data can be aggregated / stored by anyone.

All research reports

Stored in comprehensive repositories

What is the best way to find eligible studies?
Continuous update of reviews in EPPI-Reviewer

Main aim: to maintain a ‘surveillance’ of the literature as it emerges to maintain reviews up to date

New papers arrive every 2 weeks.

Machine learning models ‘learn’ the scope of each review based on included studies.

New papers are automatically added to reviews when adjudged to be sufficiently ‘close’.

Papers are included in systematic reviews in EPPI-Reviewer.

Possibility of bespoke automation in each review.
Our case study

- The EPPI-Reviewer platform contains a local copy of the OpenAlex* dataset
- It’s updated every 2 weeks
- Between 400k and 2m new records arrive each update
- We can find research using machine learning and Boolean searches

How can we use this new resource to maintain our COVID-19 map of research?

*We originally started this work using Microsoft Academic which closed at the end of 2021. OpenAlex builds on Microsoft Academic and aims to be a ‘drop-in’ replacement.
Overarching research questions

Are all the studies we need for systematic reviews indexed in OpenAlex (MAG)?

Can we efficiently identify the studies we need for systematic reviews using OpenAlex (MAG)?
COVID-19 map of research activity

- Updated weekly
- To start with, was populated via searches of Embase and PubMed
- Comprehensive
- Large team
- A lot of work!
The Octopus Study

- Aim: to undertake a cost-effectiveness study of different options for maintaining the COVID-19 map
- Lots of data in map; added detailed time-on-task data for team
- Collected data for 4 weeks of map updates (June / July 2020)
  - Compared
    - Embase / PubMed search sources
    - MAG / OpenAlex as a single search source
    - Different machine learning options
<table>
<thead>
<tr>
<th>Study Arm</th>
<th>Intervention / Comparator</th>
<th>Automation in Workflow</th>
<th>Boolean MEDLINE + Embase Searches</th>
<th>MEDLINE-Embase De-duplication</th>
<th>Automated MAG Update Search</th>
<th>Binary ML Classifier</th>
<th>Manual Screening</th>
<th>Priority Screening</th>
<th>Fixed Screening Target N = 1,500 records</th>
<th>Target Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Comparator A</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Comparator B</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Comparator C</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td>●</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Intervention A</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td>●</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Intervention B</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Intervention C</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td>●</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Intervention D</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Intervention E</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>0.95</td>
<td></td>
</tr>
</tbody>
</table>
Octopus study: results

<table>
<thead>
<tr>
<th>Gold standard included</th>
<th>Arm 1</th>
<th>Arm 2</th>
<th>Arm 3</th>
<th>Arm 4</th>
<th>Arm 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparator A</td>
<td>4378</td>
<td>4378</td>
<td>4378</td>
<td>4378</td>
<td>4378</td>
</tr>
<tr>
<td>Medline-Embase unique included</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>MAG unique included</td>
<td>743</td>
<td>743</td>
<td>743</td>
<td>743</td>
<td>743</td>
</tr>
<tr>
<td>Medline-Embase + MAG common included</td>
<td>3570</td>
<td>3570</td>
<td>3570</td>
<td>3570</td>
<td>3570</td>
</tr>
<tr>
<td>Screened-coded</td>
<td>9180</td>
<td>8722</td>
<td>6000</td>
<td>6315</td>
<td>6000</td>
</tr>
<tr>
<td>Included</td>
<td>3635</td>
<td>3455</td>
<td>2392</td>
<td>3468</td>
<td>3441</td>
</tr>
<tr>
<td>Target recall vs Source(s) included</td>
<td>1</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>Recall vs Source(s) included</td>
<td>1.00</td>
<td>0.95</td>
<td>0.66</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>Cumulative Recall vs Gold standard included</td>
<td>0.83</td>
<td>0.79</td>
<td>0.55</td>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td>Cumulative Precision vs Gold standard included</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.55</td>
<td>0.57</td>
</tr>
</tbody>
</table>

For full results, please see [https://wellcomeopenresearch.org/articles/6-210/v1](https://wellcomeopenresearch.org/articles/6-210/v1).
Further automation ...

- The map contains ~100k records manually categorized into 12 domains (diagnosis, health impacts, treatment evaluation...)
- We have plenty of training data for more refined machine learning
- We now use a BERT model in addition to the previous machine learning to put research into categories automatically
- A second model identifies research on long COVID
Full workflow

OpenAlex → About COVID-19

Irrelevant records → Can't tell

Which category does this record belong to?

Human judgement required when machine is 'unsure'

- Treatment Evaluation
- Genetics / Biology
- Transmission / Risk / Prevalence
- Social / Economic / Indirect Impacts
- Diagnosis
- Case Study - Organisation
- Case Reports (Patients)
- Treatment Development
- Mental Health Impacts
- Vaccine Development
- Long COVID
Conclusions

- We conducted a detailed study examining the costs and benefits of two different search sources together with automation options.
- Our traditional search sources had lower recall than expected.
- OpenAlex / MAG contains more research that is not indexed in Embase / PubMed than expected.
- Thanks to a lot of training data, we can also classify a large proportion of records automatically.
- With automation, we can obtain a higher recall at lower cost than our initial workflow.
- We are currently evolving the workflow with a view to maximising the efficiency gains from automation.
Thank you

James Thomas

EPPI-Centre website: http://eppi.ioe.ac.uk
Email: james.thomas@ucl.ac.uk
Twitter: James_M_Thomas
Thank You!

Evaluation:

KTDRR Webcasts:
https://ktdrr.org/training/webcasts/index.html