2nd contribution to "Systematic review generator"

Alexandru Pintea

Getting References

(20/02 - 27/02)

Getting DOI from PMID

- Since PubMed did not return DOI values for all the rows of the dataframe, I had to take the DOI values from OpenAlex though PMID
- Those newly acquired DOI values I used with CrossRef to get missing citation data
- This was done because CrossReg only takes DOI values in it's API (with OpenAlex, I got everything with PMID, no problem …)

Getting OpenAlex citations

- The only way to get citations from OpenAlex is by making multiple requests, as no actual full citations (title/authors) are referenced, just DOI links in an array …
- The only way to get those values fast (not in 80 hours, like with the first method attempted) is to ask OpenAlex with as many as 100 citations at once

Getting OpenAlex citations

- **DO THIS:** For now I limited the number of citations to 100, but I could make more than one call for all the citations for papers that have more than 100 citations
- **DO THIS:** Get permission from SCOPUS to go though their references … it's free, but I have to ask their team though an email …
- **DO THIS:** Make Git Repo for project …
- **DO THIS:** Try PMC APIs
- **DONE:** Merge back to initial CSVs …

Getting OpenAlex citations

- I got CrossRef and OpenAlex data to collect references for **14698** papers. (PubMed gave citations for **6252** papers, and how I have citations for **20950** references)
- The remaining 5848 papers that have no references might never have, as most only have a PMID and no DOI (so, I cannot use CrossRef to ger anything for them). I have also tried already to get DOIs/citations for them from OpenAlex using their PMID, but the remaining ones did NOT have either
- I have to try SCOPUS still, but first I have to send them an email asking for approval to use their citation API. Yet, most likely I won't get much from there either …

Getting OpenAlex citations

- Besides SCOPUS, I don't know where else to search … I tried PLOS too, it doesn't have much
- **OpenAlex   260 million**
- **Crossref   150 million**
- **SCOPUS   90 million**
- **ScienceDirec   21 million**
- **IEEE XPlore   6 million**
- Where are the original CSVs from? What source?

Thank you for your attention!