**Exercise database selection**

For their literature review “Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal”, Wynants et al. searched the following databases: PubMed, Embase, ArXiv, MedRxiv, and BioRxiv.

(Hint: Use the “more info” option on the databases page. MedRxiv is the medical equivalent of BioRxiv)

* + Why were these specific databases selected? (consider which types of sources and which topics these databases primarily include. Also consider the date of publication)

The aforementioned databases where selected for the research because of the authority they have for peer-reviewed biomedical literature. Some of them are freely accessible (such as PubMed), and some provide a broader pharmaceutical research and international coverage. ArXiv, MedRxiv and BioRxiv are also preprint server, offering immediate access to the latest research, which is essential when considering the fast-evolving field of COVID-19 research.

* + How do these databases complement each other to form a comprehensive search?

PubMed and Embase provide extensive peer-reviewd biomedical literature, while ArXiv, MedRxiv and BioRxiv provide early access to the latest reaserch through the preprints, allowing for greater coverage and up-to-date research.

* + Which other database(s) would you add to complement/extend this selection? Why?

Other databases may include Web of Science, which provides access to muldiple databases with cross-disciplinary research, or Cochrane Library, which specializes in systematic reviews which might provide high-quality and relevant reviews related to COVID-19.

**Exercise Search & Select – which field(s)**

Go to IEEE Xplore and carry out the searches from the table below (use command search or advanced search interface, create an account to save your search history)

* + Indicate how many results you found with each search

Search results are listed in the table below.

* + Why does search #1 find more results than search #2?

Search #1 has more results than search #2 since the former query returns all articles containing the specified words within the metadata in any order (even if not one next to the other), while the latter specifies the order and the two words must be in in order for the search to apply.

* + What leads to more results being found with search #5 than with any other search? Which sections of the articles are being searched?

Search #5 obtaines the most results since the query returns also articles that merely include the terms “machine learning” or “artificial intelligence” within the contents of the entire text.

* + Why would search #6 not retrieve all relevant/ most current results?

Search #6 would not retrieve all relevevant (or most current) entries since the query would return only the articles that have been specifically tagged or indexed under the controlled vocabulary used by IEEE.

* + What combination of searches would retrieve the optimal number of results (sensitivity + specificity)?

For optimal research results, a combination of vocabulary searches (such as IEEE Terms) with full-text searches would be the most effective.

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Field** | **Search string** | **N results** |
| 1 | All Metadata | machine learning OR artificial intelligence | **447,226** |
| 2 | All Metadata | "machine learning" OR "artificial intelligence" | **390,465** |
| 3 | Abstract | "machine learning" OR "artificial intelligence" | **140,926** |
| 4 | Title | "machine learning" OR "artificial intelligence" | **42,975** |
| 5 | Full Text | "machine learning" OR "artificial intelligence" | **578,292** |
| 6 | IEEE Terms | "machine learning" OR "artificial intelligence" | **125,984** |

**Exercise complementary search techniques**

* Find the article below in both Google Scholar and in Scopus:

Meng, L., Dong, D., Li, L., Niu, M., Bai, Y., Wang, M., ... & Tian, J. (2020). A Deep learning prognosis model help alert for COVID-19 patients at high-risk of death: a multi-center study. *IEEE Journal of Biomedical and Health Informatics*, *24*(12), 3576-3584.

* Look for documents citing this article
  + How many citing articles did you retrieve from each database?

55 citations found on IEEE, while Scopus reports only 50

* + Why might you find more articles in one database over the other?

It may have to do with the rights that a database specific has to access certain articles, and the different databases may have access to different scopes of academic journals.

* Look for related articles in Google Scholar and Scopus
* Look up the paper in Connected Papers
  + How do these three methods for finding related articles compare based on the number of results and relevance to the topic of the original article?

Looking at citing documents can produce very relevant results, since such articles build on top of the article in question, but this may not produce many diverse results and will omit many relevant articles that don’t directly cite the article in question.  
  
Looking for related articles in Google Scholar and/or Scopus can offer a broader range of related articles, nonetheless this selection is curated by algorithms and may not always be desirable or accurate.  
  
Connected Papers allows to retrieve a number of articles highly relevant to the topic in question, but might not cover a broad range of topics such as Google Scholar or Scopus.

* + What (dis)advantages do both discussed methods (citing, related) have for literature searching? Consider the difference between a scoping and systematic literature search.

In literature searching, when considering the differente between a scoping (when looking for breadth) and systematic (when looking for more depth and thoroughness) approach, analyzing cited articles can offer a scope of articles directly relevant to the original one, but may miss out on relevant research that doesn’t cite the original article directly. Looking at the related articles can instead offer a broader range of related topics, including the most current research, but may yield results that are less directly relevant to the original article.