|  |  |
| --- | --- |
| Systematic Review\* Automation Tool Journal Club  (\*and other synthesis types, too!) | |
| Date: | 18 February 2022 |
| Presenting: | Connor Forbes, Bond University |
| Tool reviewed: | Deduplicator |
| Reference: | Not yet published. Help file from the Systematic Review Accelerator webpage: <https://sr-accelerator.com/#/help/deduplicator> |
| Cost? | Free |
| Online or desktop? | Online. URL: <https://sr-accelerator.com/#/deduplicator> |
| Tool used for the following review stage(s): | Screening of references (both literature searches, and forward / backward citation) |
| Tool description from the SR Toolbox: | N/A (not listed in SR Toolbox)  Description based on the content of the help file:  The Deduplicator removes duplicates from search libraries. This tool has two main functions:  1) To deduplicate a library prior to screening (removes internal duplicates, e.g., duplicates in database search results, which arise from overlap of databases such as Embase and MEDLINE)  2) To compare two libraries and remove already screened articles (e.g., to deduplicate the forward/backward citation screen library, against the library of references from the database search). |
| Reason for reviewing the tool: | It is optimal to get the number of references to screen as low as possible without missing any relevant references. Deduplicator removes duplicate (same) references from the search results, decreasing the screening workload. |
| How the tool works: | *For a single library deduplication (search results)*   * Export your library of references from e.g., Endnote (xml format, any preferred output style); .ris and nbib are also supported * Go to the Systematic Review Accelerator’s Deduplicator page (<https://sr-accelerator.com/#/deduplicator>) * Choose one of 3 options: cautious, balanced, thorough (more information about each option is on the webpage). * Drag and drop your library (e.g. xml) * Hit upload * You will get 4 sets of results: extremely likely duplicates, highly likely duplicates, likely duplicates, and non-duplicates   Work your way through them.   * E.g. click the “extremely likely duplicates”, which will show you the versions of what Deduplicator thinks are duplicate references. * Choose which one you wish to keep by toggling the keep/discard button. * If the Deduplicator made a mistake, click the “Not Dupe” or “Split Group” buttons.   When finished:   * click file-export (top bar on the screen, left-hand side) * Choose file-export type (xml, ris, txt, nbib) * Select your preferred option for export (all references, removed duplicates, removed screened, all removed refs) * Hit Download   *For two-library deduplication (e.g. forward backward citation search results being deduplicated against the original database search results)*  Rather than the “Deduplicate” option, select the “Remove already screened and deduplicate” option and follow the steps above. |
| How the tool was tested: | So far, internal testing comparing against humans deduplicating has been conducted. |
| Test results: | N/A, see above. |
| Discussion points: | **General discussion:**   * Deduplicator currently shows title, authors, year, journal, volume & pages – what columns are shown can be adjusted in the top toolbar. * You do not have to be logged in/have an account on the Systematic Review Accelerator to use Deduplicator * We are currently planning a formal evaluation which could take a form of an RCT or a mix of simulation/human study format – to be confirmed and expressions of interest from members sought. * Potential comparators for the evaluation of the deduplication accuracy could be Endnote, Rayyan, possibly Covidence; to be approached * Cause of duplicates in search results? Overlap of content between multiple databases (e.g. Embase, Pubmed, others) * The Deduplicator performs 2 functions:  1. single library/internal deduplication (of the search results library from database searches) and 2. deduplication between 2 libraries (e.g. if you wish to deduplicate your forward/backward search results against the database search results, to avoid screening the same results twice) |