SOFTWARE FOCUS

Paperfetcher: A tool to automate handsearching and citation searching for systematic reviews

Akash Pallath | Qiyang Zhang

Abstract
Systematic reviews are vital instruments for researchers to understand broad trends in a field and synthesize evidence on the effectiveness of interventions in addressing specific issues. The quality of a systematic review depends critically on having comprehensively surveyed all relevant literature on the review topic. In addition to database searching, handsearching is an important supplementary technique that helps increase the likelihood of identifying all relevant studies in a literature search. Traditional handsearching requires reviewers to manually browse through a curated list of field-specific journals and conference proceedings to find articles relevant to the review topic. This manual process is not only time-consuming, laborious, costly, and error-prone due to human fatigue, but it also lacks replicability due to its cumbersome manual nature. To address these issues, this paper presents a free and open-source Python package and an accompanying web-app, Paperfetcher, to automate the retrieval of article metadata for handsearching. With Paperfetcher's assistance, researchers can retrieve article metadata from designated journals within a specified time frame in just a few clicks. In addition to handsearching, it also incorporates a beta version of citation searching in both forward and backward directions. Paperfetcher has an easy-to-use interface, which allows researchers to download the metadata of retrieved studies as a list of DOIs or as an RIS file to facilitate seamless import into systematic review screening software. To the best of our knowledge, Paperfetcher is the first tool to automate handsearching with high usability and a multi-disciplinary focus.

KEYWORDS
citation searching, handsearching, information retrieval, literature search, supplementary techniques, systematic reviews

Highlights
What is already known
• The current manual process of handsearching is not only time-consuming, laborious, costly, and error-prone due to human fatigue, but it also lacks replicability due to its cumbersome manual nature.
While there are several freely available tools that automate bibliographic database searching as well as forward and backward citation searching, there is a lack of automation tools for handsearching.

What is new
- This paper presents a free and open-source handsearching automation Python package and web-app, Paperfetcher, to ensure efficient and replicable article retrieval.
- In addition to handsearching, Paperfetcher includes a beta version of automated citation searching in both directions.
- The graphical web-app is available online for free and requires no programming experience to use.

Potential impact for research synthesis methods readers outside the authors’ field
- This tool can be used by researchers in any field to locate relevant studies for systematic reviews or scoping reviews.
- To the best of our knowledge, Paperfetcher is the first user-friendly tool to automate handsearching, with a multidisciplinary focus.

1 | INTRODUCTION

Systematic reviews search, synthesize and critically appraise literature on a specific topic. These are vital instruments for researchers to understand broad trends in a field and synthesize evidence on the effectiveness of interventions in solving specific problems. Performing a thorough literature search is an early and crucial step for conducting a comprehensive systematic review. The process of a literature search often begins with a field-related bibliographic database search to retrieve studies from electronic resources such as CENTRAL, ERIC, and MEDLINE, which index journals and non-journal sources. Limiting a literature search to a bibliographic database search risks missing high quality papers that were not indexed with identifiable terms, formatted as abstracts or letters, located in supplemental editions of journals, or not included in electronic databases (e.g., conference papers, preprints, or working papers). Researchers can also miss relevant studies during a bibliographic database search by failing to use an exhaustive set of search keywords or by using incorrect Boolean operators between them.

To ensure comprehensive information retrieval and transparent reporting, it is important to conduct a literature search in a systematic and exhaustive manner. Handsearching (also called ‘handsearch’, or ‘manual literature search’) is a supplementary search strategy that reviewers often adopt to identify studies missed by a bibliographic database search. It involves systematically browsing through the tables of contents of a curated list of field-specific journals, abstracts, and conference proceedings in order to gather papers relevant to the synthesis topic. Combining handsearching with a bibliographic database increases the likelihood of capturing all relevant studies and therefore underpins a solid foundation for the systematic review. For example, a validity study found that handsearching identified an additional 25% of randomized controlled trials as compared to a bibliographic database search alone. Handsearching was found to have identified between 92% and 100% of the studies included in 34 systematic reviews. This suggests the indispensable nature of handsearching among supplementary techniques.

1.1 Manual handsearching and its limitations

While handsearching traditionally involved browsing through hard copies of journals from cover to cover, technological developments have changed the practice of handsearching as researchers can now access journal articles online. One might thus argue that the term ‘handsearching’ is anachronistic, since many researchers no longer flip through journal hard copies by hand. However, despite technological advancements, handsearching is still time-consuming, laborious, and costly, and the term signals this fact. Traditionally, handsearching required reviewers or hired volunteers to manually read through the tables of contents and abstracts of tens to hundreds of journal issues, supplementary materials, and
Past literature agrees on the time and labor-intensive nature of handsearching, reporting time spent on this task ranging from an hour per journal volume\textsuperscript{15} to 185 h spent for 10 journals.\textsuperscript{8} Furthermore, while handsearching may potentially retrieve many additional unique studies, which contribute to the systematic review, in certain cases, handsearching may not retrieve any new studies.\textsuperscript{8} For example, a research team found that handsearching retrieved less than 5\% of the articles that eventually made it to their systematic review, at a rate of one useful paper for every 9 h of searching.\textsuperscript{16} Moreover, while handsearching has a higher sensitivity than a database search, identifying most of the included studies in systematic reviews, it has a much lower precision,\textsuperscript{17} that is, most studies found by handsearching are ultimately excluded. In addition, handsearching is error-prone due to human fatigue and lacks replicability due to its cumbersome manual nature.\textsuperscript{17}

Despite all its limitations, without handsearching, researchers can hardly claim their search process to be exhaustive, since a database search alone is inadequate.\textsuperscript{18} Handsearching complements a database search and serves as a cross-checking mechanism, which can potentially identify previously unseen papers. Whether handsearching is value-added can only be determined retrospectively. Therefore, the prevailing practice of handsearching urgently awaits automation to reduce the time and labor it requires. Automation of handsearching is especially beneficial given the rapid increase in studies available online and the subsequent increase in the amount of time required to retrieve papers.

\subsection*{1.2 Paperfetcher as a proposed solution to address the limitations of manual handsearching}

The current practice of handsearching requires reviewers to click through several tens to hundreds of pages of articles on journal websites to identify studies that may be relevant to the review, and then subsequently import the bibliographic metadata (which includes the article title, authors, journal, publication date, and abstract) of these identified studies into systematic review screening tools for final screening. The final screening step is performed together with articles retrieved through other search strategies. An approach for automating handsearching is to decouple the search and identification processes and make each process more efficient. The search process can be completely automated to facilitate the bulk import of bibliographic metadata of all the studies within a specified date range from multiple journals into a systematic review screening tool. Researchers can choose from several existing screening tools, which have well-structured workflows and even leverage artificial intelligence for rapid screening, to perform identification of relevant studies from the retrieved studies much more efficiently.\textsuperscript{19} Identification need not be a separate step: studies retrieved by handsearching can be combined with studies retrieved by a database search and other supplementary techniques and screened together using the same systematic review software. This greatly reduces the cost of training people. Moreover, this can help researchers systematically document and report the entire handsearching process, which can help ameliorate the issue of inadequate reporting that plagues the field of systematic reviews.\textsuperscript{20} Documenting search steps in detail improves the replicability of research syntheses and enables researchers to update synthesized results in the future.

To implement this approach for automating a handsearch, we developed a freely available Python package and an accompanying web-app—\textit{Paperfetcher}—that automatically retrieves articles from a user-selected list of journals or conference proceedings within a user-specified timeframe. \textit{Paperfetcher} returns a collated list of article titles, abstracts, and other bibliographic metadata in file formats that can easily be imported into systematic review screening tools for efficient screening (discussed in Section 2.4.) In addition to retrieving bibliographic metadata, \textit{Paperfetcher} also generates a report of the parameters used to perform the handsearching, which can assist with replicability. \textit{Paperfetcher} enables researchers to focus their energy on the screening process rather than the search process.

In addition to handsearching, \textit{Paperfetcher} also includes both forward and backward citation chasing.\textsuperscript{16,21–23} Forward citation chasing searches for articles that cite a given article of interest while backward citation chasing searches for articles cited by the article of interest. Citation chasing in both directions is a very useful supplementary search strategy.\textsuperscript{6} Researchers have found that it identifies 51\% of studies in systematic reviews.\textsuperscript{16} We implemented a beta version of citation searching in \textit{Paperfetcher} due to its important contribution to literature search.

There are a few existing tools that automate citation searching. Some of them only support uni-directional citation searching starting from a single record. Databases such as Web of Science, Scopus and Google Scholar enable forward citation searching.\textsuperscript{24} Sci-Finder is an example of a tool that supports backward citation searching.\textsuperscript{25} However, it only focuses on literature related to chemical substances and reactions. A major limitation of using tools that perform citation searching starting from a single record is that users need to engage in the time-consuming activity of performing multiple searches (one for each starting record), downloading their results, and merging them into a single dataset.
Table 1: A comparison between the process of manual handsearching and automated handsearching with Paperfetcher’s web-app

| Step           | Tasks performed by researchers in manual handsearching | Tasks performed by researchers in automated handsearching with Paperfetcher |
|----------------|-------------------------------------------------------|--------------------------------------------------------------------------|
| Planning       | Identifying relevant journals, selecting a time frame | Identifying relevant journals, selecting a time frame                     |
| Searching      | Visiting each journal’s website (or obtaining hard copies), browsing through the tables of contents of different issues within the selected timeframe | Entering the list of journals and the time frame as inputs in Paperfetcher’s web-app, clicking on the ‘Search’ button |
| Metadata retrieval and exporting | Importing each relevant article into a citation management tool such as EndNote, Zotero, Mendeley, or Paperpile to retrieve metadata. Exporting metadata for all articles from citation management software to a single file (in a file format such as RIS, EndNote XML, etc.) | Downloading the metadata for all identified articles as a single RIS file by clicking on the ‘Download results’ button. |
| Screening      | Importing the file containing the article metadata into a systematic review screening tool such as Covidence, DistillerSR, Abstrackr, or ASReview for screening based on inclusion or exclusion criteria | Importing the file containing the article metadata into a systematic review screening tool such as Covidence, DistillerSR, Abstrackr, or ASReview for screening based on inclusion or exclusion criteria |
| Reporting      | Manually documenting each step taken in the search process | Saving the report automatically generated by Paperfetcher |

A few tools support citation searching in both directions from multiple starting records. SpiderCite is one such tool that allows users to upload starting references in EndNote XML, RIS, and BibTeX formats. CitationChaser is an open source R package that performs citation searching in both directions with an accompanying Shiny app in which users can either type the DOIs of the starting records or upload a dataset of starting records in CSV or RIS formats. CoCites is a tool which performs co-citation searching in addition to forward and backward citation searching, however, it uses the National Institutes of Health’s Open Citation Collection as its data source, which focuses on biomedical literature.

However, to date, all of the above tools lack functions to automate handsearching. Paperfetcher serves to complement these tools by automating handsearching to help reviewers identify as many relevant studies as possible through supplementary search. To the best of our knowledge, Paperfetcher is the first tool which automates handsearching. In Table 1, we compared the process of manual handsearching and automated handsearching with Paperfetcher’s web-app to elaborate on Paperfetcher’s features in detail.

2 | HOW PAPERFETCHER WORKS

Paperfetcher consists of (a) a Python package that implements the core handsearching and citation searching functionalities of retrieving bibliographic information, and (b) a web application that provides an easy-to-use graphical interface for end users. Figures 1 and 2 outline the entire handsearching and citation searching workflows from end-to-end.

In the following subsections, we describe how the Python package and the web application work.

2.1 | Data sources

Paperfetcher needs to retrieve bibliographic metadata of articles from multiple journals to automate the handsearching process. It also needs to retrieve bibliographic metadata of backward and forward citations of articles for citation searching. Paperfetcher does this by querying one or more databases of bibliographic content.

Paperfetcher’s handsearching algorithm queries the database of academic content registered with Crossref. Crossref is an official digital object identifier (DOI) registration agency of the International DOI Foundation. Research suggests that Crossref is the most robust and holistic implementation of the DOI model and covers a wide range of disciplines. Among the 11 official DOI registration agencies, only Crossref, DataCite, and Multilingual European DOI Registration Agency (mEDRA) cover English materials related to scholarly and professional research content. We excluded DataCite as it only indexes research data and not research articles or reports. As mutual collaboration between mEDRA and Crossref allows DOIs registered with mEDRA to be deposited on the Crossref platform, we excluded mEDRA to avoid redundancy. Since Paperfetcher is primarily designed for English-reading researchers, we
excluded DOI agencies that contain materials of other non-English languages, such as the China National Knowledge Infrastructure, or that index non-academic content, such as the Entertainment Identifier Registry. Other non-English DOI agencies can be added in future versions of Paperfetcher once demand for multilingual research content arises.

Statistics updated in May 2021 show that Crossref indexes more than 1.3 million published and unpublished content types, such as journals, books, conference proceedings, dissertations, working papers, technical reports, and data sets. It contains bibliographic information of more than 90,000 journals and more than 80,000 conference proceedings. A comparison of different databases, including Scopus, Web of Science, Dimensions, and Crossref, concluded that Crossref is a bibliographic data source that is of significant interest for bibliometric analyses and is becoming increasingly valuable over the years. Crossref covers a larger number of documents published in journals than the widely-used Scopus database. Its wide coverage of journal articles and conference proceedings make Crossref the ideal data source for handsearching.

Paperfetcher’s citation searching algorithm queries both the Crossref database and COCI, the OpenCitations Index of Crossref open DOI-to-DOI citations. COCI is a
A comparison between the time taken to retrieve articles by automated handsearching with Paperfetcher and manual handsearching

| Journal                      | Date range                | Automated handsearching with Paperfetcher | Manual Handsearching |
|------------------------------|---------------------------|------------------------------------------|----------------------|
|                              |                           | Number of articles retrieved | Estimated time taken (min) | Number of articles retrieved | Estimated time taken (min) |
| Nature Reviews Cancer        | 2022-01-01 to 2022-05-01 | 48                        | 1                     | 48                        | 12                     |
| Review of Educational Research | 2021-01-01 to 2022-01-01 | 36                        | <1                    | 36                        | 8                      |
| Psychological Review         | 2022-01-01 to 2022-05-01 | 47                        | 1                     | 38                        | 10                     |

*Time taken by one user to navigate to the Paperfetcher app, enter handsearching parameters, start the handsearching process, wait for handsearching to finish, and download search results as an RIS file. Estimated time taken may vary according to researcher(s) experience with the tool, internet speed, and server load.

The COCI database derived from Crossref, which contains more than 1.2 billion DOI-to-DOI citation pairs of Crossref-deposited articles with open citations. The COCI database is updated periodically to add more citations.

2.2 Handsearching algorithm

Handsearching requires three inputs from the user: the International Standard Serial Number (ISSN) of the journal to fetch articles from, a date range within which to fetch articles, and an optional list of keywords to refine the search (shown in Figure 1). Paperfetcher will retrieve articles that match any of the keywords. For example, a search with the keywords ‘systematic review’ and ‘meta-analysis’ will retrieve all articles with the words ‘systematic’, ‘review’, ‘meta’, or ‘analysis’ in the article metadata. The results for each journal are ordered by the most number of matches that is, articles that contain all the keywords ‘systematic’, ‘review’, ‘meta’ and ‘analysis’ in their metadata will show up first, followed by articles that contain only some of the keywords, and finally articles that contain only one of the keywords.

Paperfetcher queries the Crossref database programatically through its Application Programming Interface (API) with these three inputs to retrieve bibliographic metadata of all articles from the journal with the specified ISSN, within the selected date range, and matching the list of keywords (if specified). This metadata contains several fields, including the article title, authors, journal, publisher, date of publication, abstract, and article keywords. The Paperfetcher web app can iterate through a list of journals, perform handsearching for each journal in the list, and merge the metadata retrieved from the search into a single dataset of handsearching results.

Table 2 compares the number of articles retrieved by Paperfetcher’s handsearching function against the number of articles retrieved by manually browsing through three journals’ websites: Nature Reviews Cancer, Review of Educational Research, and Psychological Review. Paperfetcher retrieves almost all the articles retrieved by manual handsearching from these journals. In the case of Psychological Review, Paperfetcher retrieves more articles than manual handsearching as it also retrieves issue information, editor’s acknowledgements, erratum, and commentary on previously published articles.

Although Paperfetcher supports the use of keywords for handsearching, we do not recommend using this feature when performing handsearching for systematic reviews because there is a possibility of missing relevant studies due to the use of a non-exhaustive set of keywords. The quality of database search is dependent on the choice of keywords, which is a subjective decision made by search personnel. One of the purposes of performing handsearching as a supplementary and complementary search technique is to retrieve studies missed by bibliographic database search due to incorrect or non-exhaustive keywords. Using keywords in handsearching defeats this purpose. In Table 3, we use an example to illustrate this point. Suppose a researcher wants to retrieve all meta-analyses in the journal Research Synthesis Methods published between January 1, 2022 and May 1, 2022. Automated handsearching with Paperfetcher retrieves 12 articles which contain the phrase ‘meta-analysis’ (or variations of it) in their titles, and four articles which contain the phrase ‘systematic review’ in their titles. Both sets of articles might be of interest to the researcher, as systematic reviews and meta-analysis are closely related terms with significant overlap. If the researcher was to perform automated handsearching with Paperfetcher using just the keyword ‘meta-analysis’, they would retrieve only two out of the four articles which contain the phrase ‘systematic review’ in their titles. The studies containing the phrase ‘systematic review’ which were not retrieved due to the use of keywords might have been relevant. This example shows
that using a restrictive set of keywords may result in relevant studies being missed.

Although we do not recommend using keywords when performing systematic reviews, there are cases where keywords can still be useful, such as when performing scoping reviews to identify literature in a particular research area. In such cases, we recommend using a large set of related keywords to be as exhaustive as possible. In Table 3, we show that performing handsearching with Paperfetcher using both the keywords ‘systematic review’ and ‘meta-analysis’ retrieves all the articles which contain the phrases ‘systematic review’ or ‘meta-analysis’ in their titles. It is worth noting that even if a large set of keywords is used, there is no guarantee that the set is exhaustive, and thus, there is still a risk of missing articles that were not indexed with metadata containing any of the search keywords in Crossref.

### 2.3 Citation searching algorithm

Paperfetcher requires a list of DOIs as input for both forward and backward citation searching. For backward citation searching, it queries the Crossref API to retrieve DOIs of papers cited by the input DOIs. The Crossref API only returns information about articles cited by a given paper, and not articles citing it. The COCI API, however, can return this data; therefore, Paperfetcher uses COCI for forward citation searching. In both cases, Paperfetcher compiles the retrieved DOIs into a dataset of search results, removing duplicates (if any). Figure 2 illustrates the workflow of the citation searching algorithm.

### 2.4 Data export

As shown in Figures 1 and 2, the Paperfetcher web-app allows users to export bibliographic information retrieved from handsearching or citation searching in two formats:

1. An RIS file, which contains several fields of bibliographic metadata such as article title, authors, journal, publisher, and abstract for each article retrieved from the search query. RIS is a standardized format that allows data exchange among different citation management software. RIS data exported from Paperfetcher can be seamlessly imported into systematic review screening tools such as ASReview, Covidence, DistillerSR, and EPPI-Reviewer or into citation management software such as EndNote, Mendeley, Paperpile and Zotero.

2. A plain text file consisting of the DOIs of the articles retrieved from the search, with one DOI on each line. These DOIs can be bulk-imported into citation management tools such as Zotero in order to fetch article metadata. Complete article metadata can then be subsequently exported to systematic review screening tools.

In addition to RIS and plain text, the Paperfetcher Python package can export search results into structured tabular data formats, such as Excel spreadsheets, comma-separated values (CSV) files, and pandas DataFrames. In all of these data formats, the bibliographic data fetched from the search is organized into a table in which the columns contain the article title, authors, journal, DOI, abstract, etc. Researchers can then clean, de-duplicate,
Paperfetcher

Automate handsearch for your systematic review.

What type of search do you want to perform?

Select one:

- Handsearch
- Citation search

Define your handsearch parameters.

a) Select journals to search in.

You can add multiple journals to a single handsearch. You can either search for a journal by its name using the select box on the left, or enter its ISSN in the text box on the right. To add a journal to the handsearch, click on the 'Add to search' button below your entry.

Type to search for a journal: Research Synthesis Methods, ISSN:17592887

Enter an ISSN

Add to search

The journals you selected appear below. If you wish to remove a journal from the search, click on the 'X' next to it.

Selected journals (ISSNs) to search in

- Research Synthesis Methods

b) Select a date range to fetch articles within.

Fetch from this date onwards: 2022/04/01

Fetch until this date: 2022/05/01

c) Enter search keywords [optional].

Click to expand.

d) Select output format.

You can choose from two output formats:

- A text file of DOIs (.txt): Each DOI appears on a separate line.
- An RIS citation file (.ris): Contains metadata and abstracts (if available on Crossref) for each paper.

You can directly import this file into both citation management programs such as Zotero (instructions) and systematic review screening tools such as Covidence (instructions).

How would you like to download your results?

- A text file of DOIs (.txt).
- RIS with abstracts (.ris)

Perform search

Search

FIGURE 3 Screenshot of the Paperfetcher web-app showing the handsearching interface [Colour figure can be viewed at wileyonlinelibrary.com]
Paperfetcher

Automate handsearch for your systematic review.

What type of search do you want to perform?

Select one:
- Handsearch
- Citation search

Define your citation search parameters.

a) Select the papers you want to start from.

Enter the DOIs of the papers you want to start from. You can add multiple DOIs to a single citation search. Separate DOIs with commas.

Enter comma-separated DOIs:

10.1002/ijsm.1046

b) Select type of citation search.

You can either perform backward reference chasing or forward citation chasing:
- Backward reference chasing: For each article X, find all the articles which X cites.
- Forward citation chasing: For each article X, find all the articles which cite X.

Select an option:
- Search references (backward reference chasing).
- Search citing articles (forward citation chasing)

c) Select output format.

You can choose from two output formats:
- A text file of DOIs (.txt): Each DOI appears on a separate line.
- An RIS citation file (.ris): Contains metadata for each paper.

You can directly import this file into both citation management programs such as Zotero (instructions) and systematic review screening tools such as Covidence (instructions).

How would you like to download your results?
- A text file of DOIs (.txt).
- RIS with abstracts (.ris)

Perform search

Search

FIGURE 4  Screenshot of the Paperfetcher web-app showing the citation searching interface [Colour figure can be viewed at wileyonlinelibrary.com]

and screen this tabular data exported as Excel and CSV files using various spreadsheet tools (such as Excel and Google Sheets) or statistical computing software such as R and Stata. Tabular data exported as pandas DataFrames can be efficiently processed in the Python programming language.

3  |  USING PAPERFETCHER

3.1  |  Using the web-app online

Paperfetcher's web-app, which is based on the Streamlit app framework, is available online at paperfetcher.
github.io/app. The web-app can also be run offline, as described in Section 3.4. The app has an easy-to-use graphical interface. To perform handsearching (Figure 3), users just need to fill out a few search parameters—journal names or journal ISSN, a date range to fetch articles within, optional search keywords, output format—and click on the ‘Search’ button. To perform citation searching (Figure 4), users only need to enter comma-separated DOIs of selected papers, select the type of citation searching (forward or backward), select output format, and click on the ‘Search’ button.

3.2 | Exporting data

After performing a search, Paperfetcher gives users the option to download their data either as a text file of DOIs or as an RIS file, based on the output format selection they made.

3.3 | Reporting handsearching parameters

Paperfetcher generates a report of the parameters used to perform the search and the number of papers retrieved, which is displayed to users after performing the search. Figure 5 shows a sample report for citation searching. These reports provide a mechanism for reviewers to document their search for cross-checking and reproducibility.

3.4 | Using the web-app offline

Privacy-conscious users may opt to use the Paperfetcher web-app offline. To do so, they must first install and then run the app from the source. The source code and installation instructions for the app are available online on GitHub at github.com/paperfetcher/paperfetcher-webapp. GitHub is a widely-used service to host and distribute software along with its source code, and enables us to make periodic updates to the software. Users can either download the latest version of the Paperfetcher web-app from GitHub as a compressed ZIP file or create a ‘clone’ of the app on their computers using GitHub’s desktop application or the Git command line. This ‘clone’ is a copy of the code containing its full version history, and allows users to easily pull new updates.

3.5 | Using the Python package

Users familiar with the Python programming language might prefer to directly use Paperfetcher’s Python package. The Python package has the same core functionality as the web-app; however, it can export search results to structured tabular data formats such as Excel spreadsheets, CSV files, or pandas DataFrames. Users can select which bibliographic data fields they wish to export to these tables. This is particularly useful for users who wish to perform additional data analyses (for example, removing duplicates or merging results from multiple searches) in Excel, Stata, R, or Python. Such users can install the package directly from the Python Package Index (pypi.org/project/paperfetcher). Documentation for the Python package can be found at paperfetcher.github.io/paperfetcher. Advanced Python users who are interested in modifying Paperfetcher’s code can download the source code from GitHub at github.com/paperfetcher/paperfetcher. Such users can also request new features, alert the developers of issues with the software (such as bugs or crashes), or make contributions to the source code through GitHub.

4 | DISCUSSION

The development of Paperfetcher addresses the increasing need for a more replicable, robust, and time-efficient method to conduct handsearching for systematic reviews. Although it is designed for researchers conducting systematic reviews, this tool can benefit a broader audience, including both academic and non-academic researchers who are interested in collecting information on a specific topic. Its user-friendly app requires no programming experience, and therefore, is accessible to a wide audience. In contrast to other field-focused tools, Paperfetcher can assist researchers in any field and therefore potentially has a wider user group and a larger impact. We believe that adoption of this tool can significantly reduce the time reviewers spend on a literature search and can improve transparency in synthesis reporting across disciplines.

While Paperfetcher has the potential to make significant contributions to systematic reviews, it does have a few limitations. First, there can be a time lag of 24 to 48 hours between when an article is published online and when it is retrievable using Crossref. This is the time required for DOI registration and the database update. As a result, newly published articles may not be retrieved by Paperfetcher’s handsearching and citation searching functions. Second, some publishers do not make their abstracts and/or references publicly available on Crossref. While researchers can use citation management software to retrieve missing abstracts, missing references will result in relevant studies being missed by Paperfetcher’s citation searching function. In addition, as updates to
COCI lag behind updates to Crossref, Paperfetcher's forward citation chasing function might miss newly published citing articles. Lastly, Paperfetcher may retrieve extra bibliographic entries corresponding to issue information, editor's acknowledgements, erratum, and commentary on previously published articles. Researchers using Paperfetcher will need to manually filter out this extra information.

We plan to improve upon Paperfetcher's citation searching function (which is currently a beta version) to address its limitations by incorporating additional data sources such as the Lens.org database (consisting of articles from PubMed, PubMed Central, CrossRef, Microsoft Academic Graph and CORE) and the Web of Science database, in order to increase the likelihood of identifying all references and citations. We also plan to introduce support for widely-used preprint repositories such as arXiv, bioRxiv, OSF Preprints, and MedRxiv as additional data sources for both handsearching and citation searching.

In a nutshell, Paperfetcher provides a free, easy-to-use, and efficient method to automate handsearching and citation searching for systematic reviews. To the best of our knowledge, Paperfetcher is the first user-friendly tool to automate handsearching, with a multidisciplinary focus.

AUTHOR CONTRIBUTIONS
A.P. conceived the study, developed the software, prepared visualizations, and drafted the manuscript; Q.Z. conceived the study, prepared visualizations, performed software validation, and drafted the manuscript.

ACKNOWLEDGMENTS
We thank Amanda Neitzel for insightful discussions and comments on the manuscript and for identifying bugs in early releases of the software. We thank Hunter Gehlbach for numerous constructive comments on the manuscript. In addition, we thank Susan Davis, Rohan Arcot, Aishwarya Jayabharathi for proof-reading the manuscript.

CONFLICT OF INTEREST
The authors have declared no conflict of interest.

DATA AVAILABILITY STATEMENT
Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

ORCID
Akash Pallath https://orcid.org/0000-0001-6628-6388
Qiyang Zhang https://orcid.org/0000-0001-7474-2435

REFERENCES
1. Mackenzie H, Dewey A, Drahota A, et al. Systematic reviews: what they are, why they are important, and how to get involved. 2012. Accessed May 27, 2022. https://www.semanticscholar.org/paper/Systematic-reviews%3A-what-they-are%2C-why-they-are-and-Mackenzie-Dewey/0c5d3c3f9bdf0633a778c0dd1a2939175c9f97a4
2. Cooper C, Booth A, Varley-Campbell J, Britten N, Garside R. Defining the process to literature searching in systematic reviews: a literature review of guidance and supporting studies. BMC Med Res Methodol. 2018;18(1):85. doi:10.1186/s12874-018-0545-3
3. Armstrong R, Jackson N, Doyle J, Waters E, Howes F. It's in your hands: the value of handsearching in conducting systematic reviews of public health interventions. J Public Health. 2005;27(4):388-391. doi:10.1093/pubmed/fdi056
4. Richards D. Handsearching still a valuable element of the systematic review. Evid Based Dent. 2008;9(3):85. doi:10.1038/sj.ebd.6400602
5. Kugley S, Wade A, Thomas J, et al. Searching for studies: a guide to information retrieval for Campbell systematic reviews. Campbell Syst Rev. 2017;13(1):1-73. doi:10.4073/crmg.2016.1
6. Papaioannou D, Sutton A, Carroll C, Booth A, Wong R. Literature searching for social science systematic reviews: consideration of a range of search techniques. Health Inf Libr J. 2010;27(2):114-122. doi:10.1111/j.1471-1842.2009.00863.x
7. Dickersin K, Scherer R, Lefebvre C. Systematic reviews: identifying relevant studies for systematic reviews. *BMJ*. 1994;309(6964):1286-1291. doi:10.1136/bmj.309.6964.1286

8. Glanville J, Cikalo M, Crawford F, Dozier M, McIntosh H. Handsearching did not yield additional unique FDG-PET diagnostic test accuracy studies compared with electronic searches: a preliminary investigation. *Res Synth Methods*. 2012;3(3):202-213. doi:10.1002/jrsm.1046

9. Blümle A, Antes G. Handsearching for randomized controlled clinical trials in German medical journals. *Dtsch Med Wochenschr*. 2008;133(6):230-234. doi:10.1055/s-2008-1017501

10. Salvador-Oliván JA, Marco-Cuenca G, Arquero-Avilés R. Errors in search strategies used in systematic reviews and their effects on information retrieval. *J Med Libr Assoc*. 2019;107(2):210-221. doi:10.5195/jmla.2019.567

11. Sampson M, McGowan J. Errors in search strategies were identified by type and frequency. *J Clin Epidemiol*. 2006;59(10):1057.e1-1057.e9. doi:10.1016/j.clinepi.2006.01.007

12. Higgins JP, Green S. In: Higgins JP, Green S, eds. Cochrane Handbook of Systematic Reviews of Interventions. John Wiley & Sons, Ltd; 2008.

13. Hopewell S, Clarke MJ, Lefebvre C, Scherer RW. Handsearching versus electronic searching to identify reports of randomized trials. *Cochrane Database Syst Rev*. 2007;(2):MR000001. doi:10.1002/14651858.MR000001.pub2

14. Craane B, Dijkstra PU, Stappaerts K, De Laat A. Methodological quality of a systematic review on physical therapy for temporomandibular disorders: influence of hand search and quality scales. *Clin Oral Investig*. 2012;16(1):295-303. doi:10.1007/s00784-010-0490-y

15. Moher D, Group TOC. Errors in search strategies used in systematic reviews and their effects on information retrieval. *J Clin Epidemiol*. 2006;59(10):1057.e1-1057.e9. doi:10.1016/j.clinepi.2006.01.007

16. Sampson M, McGowan J. Errors in search strategies were identified by type and frequency. *J Clin Epidemiol*. 2006;59(10):1057.e1-1057.e9. doi:10.1016/j.clinepi.2006.01.007

17. Higgins JP, Green S. In: Higgins JP, Green S, eds. Cochrane Handbook of Systematic Reviews of Interventions. John Wiley & Sons, Ltd; 2008.

18. Glanville J, Cikalo M, Crawford F, Dozier M, McIntosh H. Handsearching did not yield additional unique FDG-PET diagnostic test accuracy studies compared with electronic searches: a preliminary investigation. *Res Synth Methods*. 2012;3(3):202-213. doi:10.1002/jrsm.1046

19. Blümle A, Antes G. Handsearching for randomized controlled clinical trials in German medical journals. *Dtsch Med Wochenschr*. 2008;133(6):230-234. doi:10.1055/s-2008-1017501

20. Salvador-Oliván JA, Marco-Cuenca G, Arquero-Avilés R. Errors in search strategies used in systematic reviews and their effects on information retrieval. *J Med Libr Assoc*. 2019;107(2):210-221. doi:10.5195/jmla.2019.567
