A Literature Review of Methods, Approaches and Tools to Automate the Process of Systematic Literature Reviews

Veit Wasserfuhr

Department of Educational Sciences and Technologies, University of Twente

Master Thesis

24.06.2022

Word count: 20068

Author Note

Veit Wasserfuhr, s2586002, Faculty of Behavioral Sciences, University of Twente

Supervisor: Prof. Dr. Bernard Veldkamp, E-mail: b.p.veldkamp@utwente.nl

2nd supervisor: Dr. Maryam Amir Haeri, E-mail: m.amirhaeri@utwente.nl

Correspondence concerning this Master Thesis should be addressed to Veit Wasserfuhr Faculty of Behavioral Science, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands. Phone: +49 (0)162 6048370 E-mail: v.j.c.wasserfuhr@student.utwente.nl

Abstract

Systematic literature reviews (SLR) require a lot of resources, expertise and rigor. They are quite unpractical in their usage and are quickly outdated. Nevertheless, SLRs can bring a lot of quality output in providing an overview of newly established field or in synthesising main approaches to guide foundational research in already established fields. This thesis utilises the method of a literature review to aggregate approaches and tools that can aid in the automation of systematic literature reviews. Six dimensions have been compiled in order for the organisation of the presentation of extracted data from the research articles, which are procedural aspects, search process, screening process, selection process, data-extraction process and collaborative aspects. Problems, causes and solutions are being discussed and elaborated on. Ultimately, the thesis elicits the discrepancy between the current and the desired situation and synthesises the approaches to provide suggestions for further research in automation approaches. The guiding goal is to ask the question of how would the research community and beyond benefit, if the power of SLR could be harnessed at low cost and with high accessibility? The thesis addresses this question by attending to the role and usage of technologies in the field of systematic literature reviews.

Keywords: systematic literature review process, search process, selection process, data extraction process, Automation

Table of Contents

The Role of Automation in Systematic Literature Reviews

Literature reviews are a vital component of every scientific research process. It is the foundation that research builds upon to gain insights into what could be considered knowledge or truth. They build up the theoretical framework by informing on prior research, by defining the scope of the research and by providing context. The search and identification of all relevant literature is necessary for grounding the research in its context, to be able to hypothesise about certain aspects around a certain concept to test and to develop alternative approaches to get specific outcomes. Literature search can provide inspiration for research and can spur the testing of innovative approaches and designs. Through the use of citation and bibliography, especially in digital formats relationships between research papers and topics are generated. The literature review is a structured form of inquiry, nevertheless exhaustiveness and rigor are lacking in many literature reviews. Therefore, the field of Systematic Literature Reviews (SLR) has been developed. They are a powerful form of research that take an essential role in scientific research. They are seen as the highest quality of research articles after meta-analyses in the hierarchy of evidence (Cook et al., 1997). In order to adhere to the high standards of scientific discourse and to ensure the quality of literature reviews, SLRs are highly structured and focused on methodology. Nevertheless, they are a tedious, complex and time-consuming endeavour in every literature research process, with many of its tasks being done manually in the solitude of one research project (Koukal et al., 2014; Webster and Watson, 2002; Felizardo et al., 2011; Goldfarb-Tarrant et al., 2020).

Problems of Systematic Literature Reviews

The main problem of systematic literature reviews is that they usually require between 6 months and 2 years (Khangura et al., 2012) and completing a systematic literature review takes approximately an average of 1139 hours (Allen & Olkin, 1999). Furthermore, approximately 70% of the time spent working on the review is used to conduct the search and selection process of potentially relevant literature. A formula derived by Michelson & Reuter (2019) estimates that a typical SLR in a clinical setting costs approximately $141,194.80. Felizardo et al. (2011) describe that the selection of primary studies in SLRs can be arduous, especially when it involves a large volume of possibly relevant studies and Webster and Watson (2002) describe that literature reviews are more time-consuming and have fewer research outlets than research articles. One of the reasons for that are their high standards and quality control (Kitchenham & Charter, 2007; Leidner, 2018).

Causes of Problems in Systemic Literature Reviews

The main causes of the problems in the search process are limited meta-data, semantic integration in search engines, annotation and labelling of research papers, the lack of sharing procedural and labelling information between research projects and the lack of cooperation and collaboration between researchers, culminating in researchers carrying out similar tasks repeatedly. Search engines have helped the search process, however, "they too often create a large basket of articles that must be read to detect those studies pertinent to the matter of interest" (Watson & Webster, 2020, p.9). Therefore, the "fundamental problem is that knowledge is not encoded, and scholars must rely on the methods of their forebears (reading or scanning many papers) to take a step forward" (Watson & Webster, 2020, p.9). Furthermore, splintered databases and limited access to scientific literature by paywalls increase the complexity of the search and selection process. Another problem is the lack of knowledge of and a lack of familiarity with the procedure of conducting a systematic literature review, which has been described by Webster and Watson (2002, p.14) as "many Information Systems (IS) scholars are not familiar with the structure and format of reviews." Moreover, van Altena et al. (2018) found that even readily available automation tools are currently not widely used among participants. The authors concluded that tools are usually used by recommendations of peers and that licensing, steep learning curves, lack of support and a mismatch to the workflow are relevant barriers to the uptake (van Altena et al., 2018).

Solutions to Problems in Systematic Literature Reviews

Several solutions to tackle problems in SLR exist in the literature which are discussed, elaborated and synthesised in the literature search section. A multitude of possible approaches are described and discussed, like visual text mining techniques (Felizardo et al., 2011), crowdsourcing (Nama et al., 2019; Mortensen et al., 2017), natural language processing (Atanassova et al., 2019), provision of guidelines (Bai et al. 2019; vom Brocke, 2015; Webster & Watson, 2002) or a combination of human expert annotations and machine learning (Goldfarb-Tarrant et al., 2020). Moreover, a multitude of digital tools have been developed to support the SLR process, like for example Elicit, connectedpapers, Notion, Zotero, Mendeley, ASReview, or Covidence.

The solutions that have been elicited can be grouped in the following topics and areas: Types of reviews (Typology), Procedural aspects, Research management solutions, Guidelines, Frameworks, Protocols, Checklists and Design principles, Domain specific issues, Semantic analysis approaches, Indexing mechanisms, Bibliometrics and Citation

networks, Graph Theory, Concept Mapping, Machine learning approaches, Text mining, Crowdsourcing, Summaries, Classification and Annotation, Collaboration tools, Crowdfunding and Channels for distribution.

These approaches and more are utilised in order to answer the following research question: *What methods, approaches and tools can be used to inform the design of improving the effectiveness and efficiency of the search, selection and data extraction process in systematic literature reviews?*

Project Plan

This project will explore the field of systematic literature reviews, specifically focussing on the search, selection and data extraction process of relevant literature. In the context of an aggregation and evaluation project, it is aiming for a review of methods currently being used in the field of SLR and of proposed novel and promising methods in the field. Further, it is trying to propose recommendations for conducting the process of SLR, especially regarding improvements in the efficiency of each step in systematic literature reviews based on a synthesis of existing approaches. It is a concept-centric approach with the unit of analysis being the six dimensions identified in the literature search section, as proposed by Webster & Watson (2002) which aims at finding solutions to the problems in the search, selection and data extraction process, as well as in the procedural and collaborative aspects, in order to create an overview of the possibilities that SLRs can provide for the scientific discourse. It is of key importance to reduce the amount of time and resources spend on searching and selecting relevant literature, as this would reduce the costs and efforts of the SLR process. Ultimately, reducing hurdles in the process of SLR would improve the willingness to conduct a systematic literature review and would improve the attitude of researchers taking on the tedious process (Goldfarb-Tarrant et al., 2020).

Desired Situation

The desired situation is that researchers are able to define their research question and the corresponding inclusion and exclusion criteria, plug them into a software tool and consequently receive a complete and comprehensive list of relevant literature and a visualisation of the relation between the papers in that topic in order for the researchers to be able to immediately start working on answering their research question. Furthermore, it would be very helpful if the relevant parts of the papers are highlighted which reduces the search time within a paper and helps in the labelling process.

Theoretical Framework

The first question to ask is what do we get from academic publications? They can describe concepts, hypotheses and theories and provide an overview of the findings and conclusions that can be drawn from the research. Each publication has to include the following components: The research question, concepts, hypotheses and theories, research designs, data collection methods, scales and operationalisations, data analysis methods, findings and conclusions. Optimally, they also include a future outlook on the field of research. Furthermore, the scientific progress relies upon scholars synthesising existing work to lay a foundation for future research (Watson and Webster, 2020).

Types of Reviews and Typology

The general aim of a systematic literature review is to find, assess, analyse and report on literature that is relevant for the proposed research question, which is based on the problem statement. The general outline of a systematic literature review starts with the formulation of a research question, followed by designing the study and creating the project plan. Next, the search for related literature is carried out, which can be done using databases and search engines. Afterwards, the selected literature is screened, and the quality of the study is appraised. In this process the articles are reviewed for relevance by applying the defined inclusion and exclusion criteria. This process is followed by the extraction of the relevant information from the articles, which are aggregated and connected to answer the research question. In order to report on the research, the findings and the process itself the systematic literature review document is created. This document is an academic article, which is following the scientific writing standards (Xiao & Watson, 2019).

Two types of reviews exist. First, “authors could deal with a mature topic where an accumulated body of research exists that needs analysis and synthesis.” (Webster and Watson, 2002, p.14). Second, authors could tackle an emerging issue that would benefit from exposure to potential theoretical foundations. The author's contribution would arise from the fresh theoretical foundations proposed in developing a conceptual mode” (Webster and Watson, 2002, p.14). Webster & Watson (2002, p.13) state that an effective review creates “a firm foundation for advancing knowledge.” Webster & Watson (2002, p.18) believe that “sense-making is enhanced when a review is logically structured around the topic's central ideas and makes good use of tables and figures to convey economically the key findings and relationships.” All review types should be ‘systematic’, in the sense that all research is expected to follow some ‘system’ of inquiry (Booth et al., 2016). There are two points in a

scholar's life that necessitate a literature review. First, if substantial progress on a stream of research has been achieved and the researcher is well positioned to tell their colleagues about what they have learned and second, scholars who have completed a review prior to embarking on a project (Webster and Watson, 2002).

Xiao and Watson (2019) summarise the typology of literature reviews. They describe the procedures for conducting a review and provide tips on how to improve the quality and rigor of literature reviews in planning education and research. Rigorous systematic literature reviews have two important aspects. First, the followed steps are clearly documented and second, the review makes valuable scholarly contribution (Okoli, 2015). Furthermore, vom Brocke et al. (2009) express that explicitness in documenting the methodological details of a review by publishing the documentation is necessary for proving the conduction of a rigorous approach. Moreover, Paré et al. (2015) identified a typology of nine ideal review profiles, which are theoretical reviews, narrative reviews, meta-analyses, descriptive reviews, hybrid reviews, critical reviews, and scoping reviews.

Literature reviews can take two forms, which are background reviews and stand-alone reviews. Background reviews are commonly used as justification for decisions made in research design, provide theoretical context, or identify a gap in the literature the study intends to fill (Templier & Paré, 2015; Levy & Ellis, 2006). In contrast, stand-alone reviews attempt to make sense of a body of existing literature through the aggregation, interpretation, explanation, or integration of existing research (Rousseau et al., 2008). Okoli (2012) provide detailed guidelines to writing a high-quality theory-mining review and distinguishes between three types of reviews that seek to build and contribute to theory, which are theory-landscaping, theory-contending and theory-testing reviews.

Systematic reviews are integrative articles. Other important examples of integrative articles are practice guidelines and clinical decision analyses. They can help practitioners keep up to date with the overwhelming volume of for example, medical literature. Ultimately, systematic literature reviews can help ground for example, clinical decisions in research evidence (Cook et al. 1997).

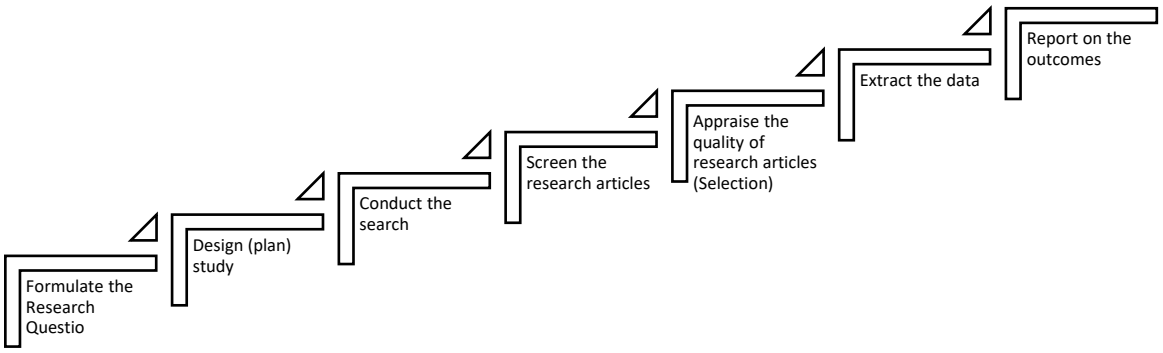
Process of Systematic Literature Reviews

In order to grapple with the ideas and methods for improvement of the SLR process it is necessary to understand and have a good overview of the full process. Three major stages are involved in a successful review (Kitchenham & Charters, 2007). The first stage is

planning, where the need for a review is identified and a review protocol is developed. The second stage is conducting the review, which is about the identification and selection of primary studies, which are reviewed by extracting, analysing and synthesising the data. The third stage is reporting the review, in which the dissemination of the findings from the literature review are written in a report (Kitchenham & Charters, 2007). Synthesising from the most commonly used SLR processes, 7 major steps could be identified, which build on their previous steps, according to a synthesis of Webster and Watson (2002), Xia and Watson (2019), Petticrew and Roberts (2006) and (Kitchenham & Charters, 2007) and are outlined in Figure 1:

Figure 1

Synthesised Steps of SLR from Literature



Note. This figure describes the synthesised steps for SLR which has been adapted from Webster and Watson (2002), Xia and Watson (2019), Petticrew and Roberts (2006) and (Kitchenham & Charters, 2007).

The following list describes the activities that need to be conducted in each of the seven steps outlined in Figure 1.

1. Research Question: Formulate an appropriate and answerable research question

2. Design (plan) study: Define scope of the study and define what relevant literature is by describing the inclusion and exclusion criteria.

3. Search: Using keywords / search terms in databases, which are derived by the research question

4. Screening: Using the title, abstract and meta-data to determine whether the paper can inform the review and answer the research question.

5. Quality appraisal: The quality and eligibility of identified relevant research articles are evaluated. This step results in the selection or discarding of the research article.

6. Data Extraction: Selected papers are examined in full-text and relevant aspects are extracted to be used to inform the review and answer the research question. These data can be of multi-media nature. The findings are analysed and synthesised in order to arrange them in a coherent and presentable structure.

7. Report outcomes: Aggregation and synthesis of information which resulted from the previous step in an accurate, well-structured and digestible format, describing each step in the process and providing an overview of literature that has been engaged in with.

The importance of the research question determines all other aspects of a review. The question is the aim of the review and the basis for the selection of a fitting method for the review.

The guiding definition of the International Collaboration for the Automation of Systematic Reviews (ICASR) are the “Vienna Principles”, which states that high standards need to be adhered to SLR for planning, reporting, conducting and updating rigorous reviews. In the commentary by O’Connor et al. (2018) the authors refer to a tool as a software application with a user interface that fully or partially automates a task conducted by systematic reviewers.

Purpose of Systematic Literature Reviews

The purpose of a SLR is to increase the understanding of a topic by optimally considering all available information and to consequently combine evidence in a meaningful way. A systematic review attempts to identify and analyse the best available evidence in order to answer a research question by building on the findings and reasoning of existing literature. They are a condensed overview of multiple research papers, each of which have their specific content, focal points and methodologies. SLRs further provide a foundation of knowledge on a topic, identifies inconsistencies and gaps in the literature and outlines relationships between papers and topics. Evidences, which differ in their relevance to the

proposed research question, from different research methods and sources are aggregated and synthesised. Therefore, the reliability and the precision of each step in the process are of high importance. However, combining disparate evidence meaningfully is a challenge. Nevertheless, analysing such evidence with a systematic approach reduces observer bias (Glade, 2008).

A literature review is an essential feature of academic research, due to the fact that they establish the foundation of academic inquires. Fink (2020) defines a research literature review as a systematic, explicit, and reproducible method for identifying, evaluating, and synthesizing the existing body of completed and recorded work produced by researchers, scholars, and practitioners" (Fink, 2020). Systematic reviews assemble, critically appraise, and synthesize the results of primary investigations addressing a specific topic or problem, that are prepared using strategies that limit bias and random error (Cook et al. 1997). Fundamentally, knowledge advancement must be built on prior existing work. By reviewing relevant literature, we understand the breadth and depth of the existing body of work and identify gaps to explore. By summarizing, analysing and synthesizing a group of related literature, we can test a specific hypothesis and develop new theories. It is also possible to evaluate the validity and quality of existing work against criteria to reveal weaknesses, inconsistencies, and contradictions (Paré et al. 2015). Literature reviews, as scientific inquiries, should be valid, reliable, and repeatable (Xiao & Watson, 2019).

What can a systematic literature review ultimately do? A review should "identify critical knowledge gaps and thus motivate researchers to close this breach. That is, writing a review not only requires an examination of past research, but means making a chart for future research" (Webster and Watson, 2002, p.19). It should also highlight "the discrepancy between what we know and what we need to know, which alerts other scholars to opportunities for a key contribution" (Webster and Watson, 2002, p.19). Furthermore, it should "extend current theories or developing new theories will create directions for future research, which is the most important part of a review and generally needs the most resources" (Webster and Watson, 2002, p.19). The reasoning for propositions may come from three main sources: "theoretical explanations for 'why', past empirical findings and practice or experience. The why or logical reasoning is the most important component of the explanation" (Webster and Watson, 2002, p.19). In conclusion, a review paper embodies the 'state of the field'. As such, it "represents a benchmark for others conducting future research in your area" (Webster and Watson, 2002, p.20). As with any research, literature reviews do

not present an end in themselves but are conducted to serve a certain purpose (Okoli & Schabram, 2010). A literature review is both an outcome and a method and from a methodological perspective, literature reviews are guided by research problems, which provide their justification.

Methodology

The following sections describe the steps that have been taken in order to answer the research question. It is a detailed description of the tasks that have been carried out to conduct a rigorous and comprehensive overview of the current state of the literature and explicates the reasoning behind the decisions that have been made during the process of conducting the review.

Procedure

This review commenced with the construction of a pilot search, which included a general search for potentially relevant articles by using the keywords from the research question and using databases such as Web of Science, Scopus and the University of Twente library. The following search terms were used to find relevant publications to answer the research question. These are aiming to find improvement, optimisation and automation approaches in conducting systematic literature reviews. ‘Improvements ∨ Optimisation in conducting Systematic Literature Research’; ‘Improvements of the Systematic Literature Reviews Process ∧ Search ∧ Selection’; ‘Optimisation of the Systematic Literature Review Process ∧ Search ∧ Selection’; ‘Automation of Systematic Literature Reviews’ (Core search term).

The research papers that have been identified through the screening of the titles and abstracts in the search results have been used to inform the theoretical foundations of the search and provided a glimpse into the emerging field of research. Nevertheless, instead of systematically screening all the tens of thousands results from the keyword search, the approach of snowballing has been used. A pool of 19 articles informed the pilot search and was the basis for the extended literature search. Furthermore, the insights from these publications have been used to inform the methodological procedure.

The Systematic Literature Review process suggested by Petticrew & Roberts (2006) has been followed in order to conduct the literature review part of this thesis. This stepwise process encompasses formulating the research questions, defining the search terms, selecting the databases, conducting the literature search, formulating inclusion and exclusion criteria

and applying these to selected relevant literature and finally the extraction of data. In order to report on these findings a reproducible method has to be used. Staples and Niazi (2007) emphasize the need to keep a record of what happens during the conducting of the review and they point out the need to report deviations from the protocol. Therefore, the issue that arose during the search process is described in detail. The main problem was the sheer number of publications that resulted from the search in different databases using keywords. Most of the publications found in the search process were either about improving, optimising or automating some processes in a field of research which was not related to SLR or they were SLRs about a specific topic or a specific field which are not relevant in answering the research question. Therefore, in order to find all relevant research papers to answer the research question of this thesis, alternative methods to screening more than 200.000 articles had to be found, due to the fact that it would otherwise not be feasible to conduct this review. Consequently, the only viable alternative was to use already included and selected relevant research papers to find related papers. The most prominent alternative method for that purpose is to use the snowballing technique, which can either be done in a forward or a backward fashion (Webster & Watson, 2002).

One way of conducting snowballing in this thesis was to identify and search leading journals and conference proceedings about the research topic. Another way was the backward search by reviewing the reference lists of already identified relevant publications and iterate until no new papers could be identified. The forward snowballing method identifies relevant publications by finding articles that cite already identified articles. During the search process a recently developed tool called ‘Connected Papers’ has been identified to discover relevant research papers. This tool enables the search of one paper by name, title or other meta-data to generate a graphical representation in form of a mind-map, that shows relationships and linkages between and across papers. All papers are ordered according to their respective relationships represented by their cluster-membership, hence similar papers have strong connecting lines and cluster together. The node size represents the number of citations and the node colour is the publishing year. On top of that, already identified papers were searched in the Wiley library, where 5 related articles are shown which could be screened for relevance. The Semantic Scholar Website incorporates a feature that shows a list of recommended papers, provides a button for a list with the references of the paper and a button for a list of citations of the paper. Moreover, there is a button which shows related research papers.

Another approach taken to find relevant literature was to discuss the topic with peers and colleagues, which brought forward interesting articles that could be used for answering the research question and used for snowballing purposes.

It was a deliberate choice to use this exploratory search approach on top of the initial start of the search process commencing with the conventional systematic literature review process. These alternative methods were able to find an extensive number of relevant research papers and additional papers that would have been missed by the search terms, hence the keywords that have been used. It could be called a systematic snowballing approach. By utilising this method, most of the relevant papers should have been uncovered. Ultimately, the aforementioned procedure resulted in having a corpus of 90 research publications which can be used to answer the research question. These selected research articles discussing the automation of steps in the SLR process can be grouped into six categories, focusing on the search process, the screening process, the selection process, the data extraction process, the collaborative process and procedural aspects. These categories will be elaborated on in the literature search section.

Data Management

In order to manage the selected research articles from search engines and journals, a folder incorporating the relevant literature in pdf form has been created. Relevant sections from the research paper are highlighted using a colour scheme to differentiate between definitely relevant, good to know, technical aspects and general aspects. The deemed definitely relevant information are extracted from the scientific publications into a text processing application. A citation system called Zotero has been used to manage references in the bibliography with relative ease. The outcomes for which data will be sought is based on answering the research question. The main outcomes are information regarding the search, selection, screening and data extraction process, as well as procedural and collaborative aspects in the SLR process. The information extracted from the papers are summaries of the respective paper that aid in explaining concepts and approaches that can be utilised for the development of a SLR.

Databases and Search Terms

The main databases utilised for the literature search are outlined in Table 1. The table describes the purpose for using the database and the number of results from each search.

Table 1

Databases, Purpose and Hits of Keyword Search

Database	Purpose	Hits
Utwente library (Ut.on.worldcat.org)	screening	4,113 (04.10.2021)
IEEExplore website	screening	128 (14.10.2021)
IS Systems Journals	screening	32 Volumes (12.10.2021)
Sciencedirect.com	Screening using advanced search criteria	21,934 (04.10.2021)
Onlinelibrary.wiley.com	Screening and related research paper feature	76.196 (04.10.2021)
Connectedpapers.com	Snowballing from existing relevant research papers	Mind-map of related research papers 40 papers (20.10.2021)
Google Scholar	To get access to specific research papers that were not available to download from other databases	771.000 (04.10.2021)

Note. This table demonstrates the databases that have been searched, the purpose for use and the number of results that could be retrieved for the search phrase: Automation of Systematic Literature Reviews.

Selection Process

The titles and abstract of research articles from the searches were screened by applying the following inclusion / eligibility criteria outlined in Table 2. Each publication was reviewed and a decision was made whether the criteria is present or not present. If at least 2 of the criteria have been met, the publication proceeded into the next round.

Table 2

Inclusion and Eligibility Criteria

Inclusion and eligibility criteria	Yes / No

The paper was published in a scientific, peer reviewed journal, was a dissertation, a meta-analysis, a systematic literature review, a book (chapter) or a conference proceeding.
The paper describes the structure and process of Systematic Literature Reviews.
The paper describes one or more validated and implementable tools that improve / optimise the search process of Systematic Literature Reviews.
The paper describes one or more tools that improve the selection process of Systematic Literature Reviews.
The paper describes the conceptual understanding of a research paper.
The paper describes the topic in the context of Systematic Literature Reviews on a meta-level.

Note. This table outlines the inclusion criteria to aid in the decision-making process of finding relevant research publications for the literature search section.

The decision to decision to use the eligibility criteria has been based on the aim to include high quality publications about the structure, processes, tools, concepts and context of Systematic Literature Reviews.

Data Extraction and Analysis

In order to assure the quality of each included research paper, the quality appraisal criteria have been developed. Each publication is screened for the quality appraisal criteria and is rated either as being high, moderate or low.

Each of the selected publications' full-text was read and relevant results were recorded using a data extraction form. The data extraction form was generated over multiple iterations to ensure usability and consistency in the data extraction procedure. The data extraction form consists of categories that are generally and specifically relevant information that can be extracted from research articles, which can be used to differentiate between research articles. Furthermore, this enables the ability to identify categorical membership of specific parts of a research paper. This process has been iterated along the data extraction

process. Table 3 tries to build up a comprehensive list of possible properties that a research article can have and as a general description lists extractable information from research articles. Moreover, Table 3 outlines potentially relevant information that can be included in a research publication and is an overarching general data extraction form. The table was developed by synthesising the literature and by updating it based on the information found in research publications.

Table 3

Meta-data of Selected Relevant Research Papers Framework

Aspect	Content
General	
Title	
Publishing Year	
Author(s)	
Subject	
Keywords	
Type of Research Paper (Research Approach)	
Quality appraisal	
Journal	
Within paper	
Topic and Theme	
Structure of paper (headings) / Framework	
Levels of Analysis	
Argumentation structure	
Logic	
Theoretical	
Analytical	
Empirical	
Methodology	
Hypotheses	

Content in table format
Quantitative data
Grafical content
(Pseudo)-Code
Results
Theoretical Approach
Conceptual Approach
Arguments (new / cited)
Data Sources
Reference list (classify in sub-categories)
Practical example
Context
Contextual variables

Note. Meta-data and more of selected relevant research papers framework. Outline of the aspects that publications can incorporate. The levels of detail that an analysis of a publication can have.

All of the studies were blindly double-coded to confirm the reliability of the inclusion of research articles and to reduce potential biases. This process resulted in an agreement of 93.42%, which is an inter-rater-reliability score (Cohen's Kappa) of 0.645 has been achieved. After all differences were discussed, 5 out of 95 papers have been excluded and an agreement after discussion of 100% has been achieved.

Quality Appraisal

The strength of the body of evidence will be assessed using the Quality appraisal criteria in Table 4. Only papers that score moderate or higher in the criteria 1-4 and low in criteria 5 and 6 are included in the review. The assessment of the risk of bias at the study or outcome level is addressed by incorporating different viewpoints on the same topic, hence by not excluding research articles based on specific viewpoints that support the initial tendency for answering the research question. During the data synthesis all relevant information is being extracted from the research articles. Meta-biases in the publication of articles about SLRs have not been identified. However, it is not possible to completely rule out the

existence of meta-biases. Selected reporting might be caused by the publication of implementable concepts or guidelines.

Table 4

Quality Appraisal Criteria

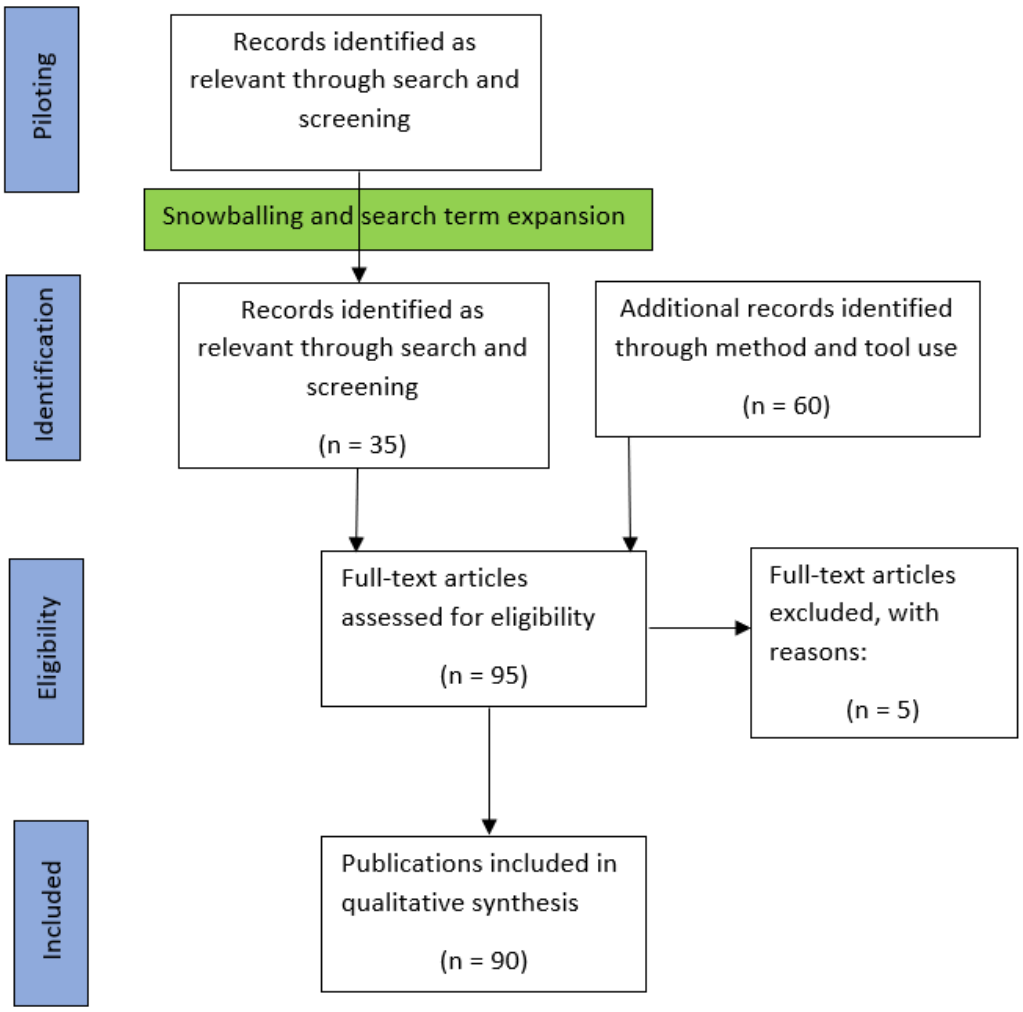
Criteria	Rate (high/moderate/low)
1. Type of Article (Journals, scientific literature, Systematic Literature Review, Meta-Analysis)	
2. Quality of reporting	
3. Reputation of authors	
4. Implement ability of proposed tool	
5. Potential biases that influence results	
6. Authors' conflict of interest	

Note. Quality Appraisal Criteria used to identify if the publication fulfils the criteria to be included in the review.

In order to provide an overview of the process of the Thesis and to provide a detailed description of the tasks that have been carried out a flow chart has been created. Figure 2 outlines the phases of the review and describes the number of publications that are relevant.

Figure 2

Overview of the Process of this Thesis using a Flow Chart



Note. Overview of the literature search process of this review as a flow chart.

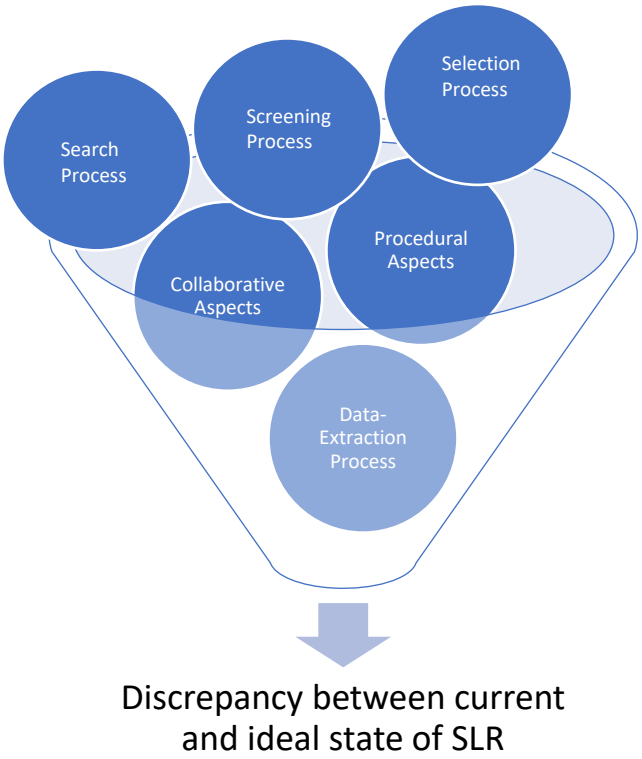
Literature Search

The following sections are structured by opening with a description of the individual aspects of each step in the process, as elaborated in the method section. The selected research articles are categorised into 6 areas of the SLR process which are subject to automation in current literature, which are procedural aspects, the search process, the screening process, the selection process, the data extraction process and the collaborative aspects. Each topic is summarised, whereby an overview of the topic and methods in the literature is provided at the beginning of each section, followed by the introduction to the topic is provided. Each method

is summarised and a detailed description of the theory and method in scope are provided. An overview of the literature search process is provided in Figure 3.

Figure 3

Outline of the Literature Search Section



Note. Funnel outlining the flow of the literature search section.

Automating the steps in the systematic review process provides opportunities for conducting systematic reviews faster, especially in health care, with fewer resources, to produce more reviews, to answer more clinical questions, to keep them up to date and to provide evidence-based answer to clinicians (Tsafnat et al., 2014).

Felizardo and Carver (2020) aggregated the most common challenges for SLRs from the literature. These are developing the review protocol, searching for evidence, selecting relevant studies, extracting data and synthesising the evidence. The authors state that "the overall goal of an SLR is to synthesize information presented in a disparate set of studies and present the findings to the reader in an easily understandable fashion" (Felizardo & Carver, 2020, p. 5).

Procedural Aspects

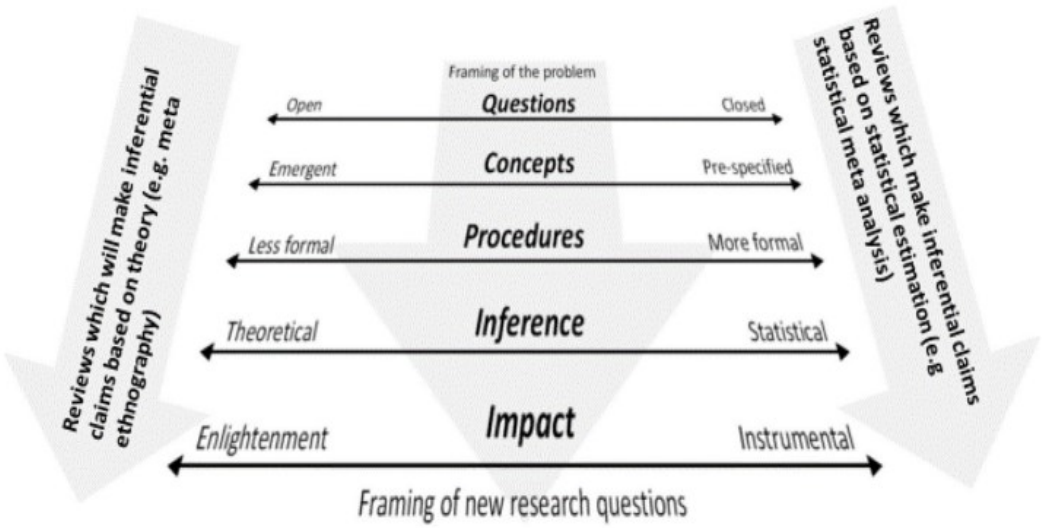
Procedural aspects are concerned with opportunities and issues regarding the general process of SLR. These include the development of guidelines (Bai et al., 2019; Okoli, 2015; Wolfswinkel et al., 2013; Xiao & Watson, 2019) the proposition of a general framework (da Silva Júnior & Dutra, 2021; Webster & Watson, 2002; Kitchenham & Charters, 2007; Petticrew & Roberts, 2006), the creation of a SLR Specification manual (Glade, 2008), the provision of an overview of automatable processes (Beller et al., 2018; Tsafnat et al., 2013; Tsafnat et al., 2014; Van Dinter et al., 2021), the investigation of the current level of uptake of automation tools (van Altena et al., 2018), the application of the process to specific domains like health care (Cook et al., 1997) and software engineering (Brereton et al., 2006), the assessment of labour costs (Michelson & Reuter, 2019), the introduction of alternative approaches, for example the rapid review approach (Khangura et al., 2012), the use of research management systems (Naak et al., 2020; Covidence), the support of management of SLRs (Bowes et al., 2012) and the encoding core knowledge of research articles (Watson & Webster, 2020).

Research Question

The importance of the research question determines all other aspects of a review. The question is the aim of the review and the basis for the selection of a fitting method for the review. Research questions can be divided into different types concerning the collection of specific data based on concepts and theories or are focusing on developing and generating concepts. Figure 4 outlines the differences between inferences based on theory and statistical estimation regarding questions, concepts, procedures, inference and impact to frame a research question.

Figure 4

Dimensions of Difference in Approaches to Synthesis



Note. Dimensions of difference in approaches to synthesis adapted from Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews* (2nd edition). SAGE.

Research Management Systems

Research Management systems can be a solution to organise the review process and to improve and simplify and increase error proneness of the whole systematic literature review endeavour. Several Research Management systems have been developed, for example Papyres (Naak et al., 2008), Buhos, Covidence (http://www.covidence.org, retrieved 15.10.2021) ASReview (van de Schoot et al., 2021), (https://www.asreview.nl, retrieved 15.10.201), Mendeley (https://www.mendeley.com, retrieved 28.11.2021), Zotero (https://www.zotero.org, retrieved 28.11.2021), Obsidian (https://www.obsidian.md, retrieved 16.11.2021), Calibre (https://www.calibre-ebook.com, retrieved 16.11.2021), Rayyan (https://www.rayyan.ai, retrieved 25.11.2021), Eppi-reviewer (https://www.eppi.ioe.ac.uk/cms, retrieved 23.11.2021), SESRA (Molleri & Benitti, 2013) (http://www.sesra.net/index/index, retrieved 19.11.2021) and many others (http://www.systematicreviewtools.com, retrieved 14.11.2021).

Naak et al. (2008) explicated the Research Paper Management System called Papyres. It was one of the first tools to synthesise the features of several narrow management systems. Papyres combines functionalities from bibliography management systems with paper

recommendations techniques and an enterprise content management (ECM) system to offer a complete environment for managing research literature and for sharing of knowledge. (Naak et al., 2008). Papyres provides functionalities like status, comments, review, tags, forum, RSS (really simple Syndication) and document organisation features such as folders, subfolders, access control, classification and custom links.

Overview of Automatable Processes

Several papers are highlighting the need for automation of SLR and describing systematic review automation technologies (Beller et al., 2018; Tsafnat et al., 2013; Tsafnat et al., 2014; van Dinter et al., 2021; da Silva Júnior & Dutra, 2021). The authors mention examples of tools used for the automation of evidence synthesis tasks. “Much research on integrated automatic review systems is still needed to execute a series of search, appraisal, information extraction, summarization, and report generation algorithms on all the data available” (Tsafnat et al., 2014, p.12) and propose a systematic review protocol which can be modified, updated, corrected and then redistributed so that such amendments are reflected in new systematic reviews based on the review protocol. They suggest that “Semi-automated decision support systems will advance the end goal of completely autonomous systematic review systems” (Tsafnat et al., 2014, p.12).

Beller et al. (2018) cover the need for improvements in the efficiency of systematic review tasks. The authors provide an overview of automation tools for specific processes and highlight the need for collaboration and varied skills in the SLR field. Collaboration between researchers need to be supported by integrated software solutions. The ICASR group devised the Vienna principles, which discuss the need for approaches of the automation of SLRs. These principles state that high standards need to be adhered to for planning, reporting, conducting and updating rigorous reviews. Furthermore, automation tools and techniques need to be evaluated using replicable methods and the results and data need to be reported. An area that still needs to be addressed is the integration of SRs into the knowledge translation process.

Proposition of a General Framework

Van Dinter et al. (2021) aim to collect and synthesize the studies that focus on the automation of SLR. The authors state that all processes can be automated. However, it is yet difficult to automate the reporting process of the review, as well as to identify the need for a review, as both processes require human creativity and insight. The authors identify and

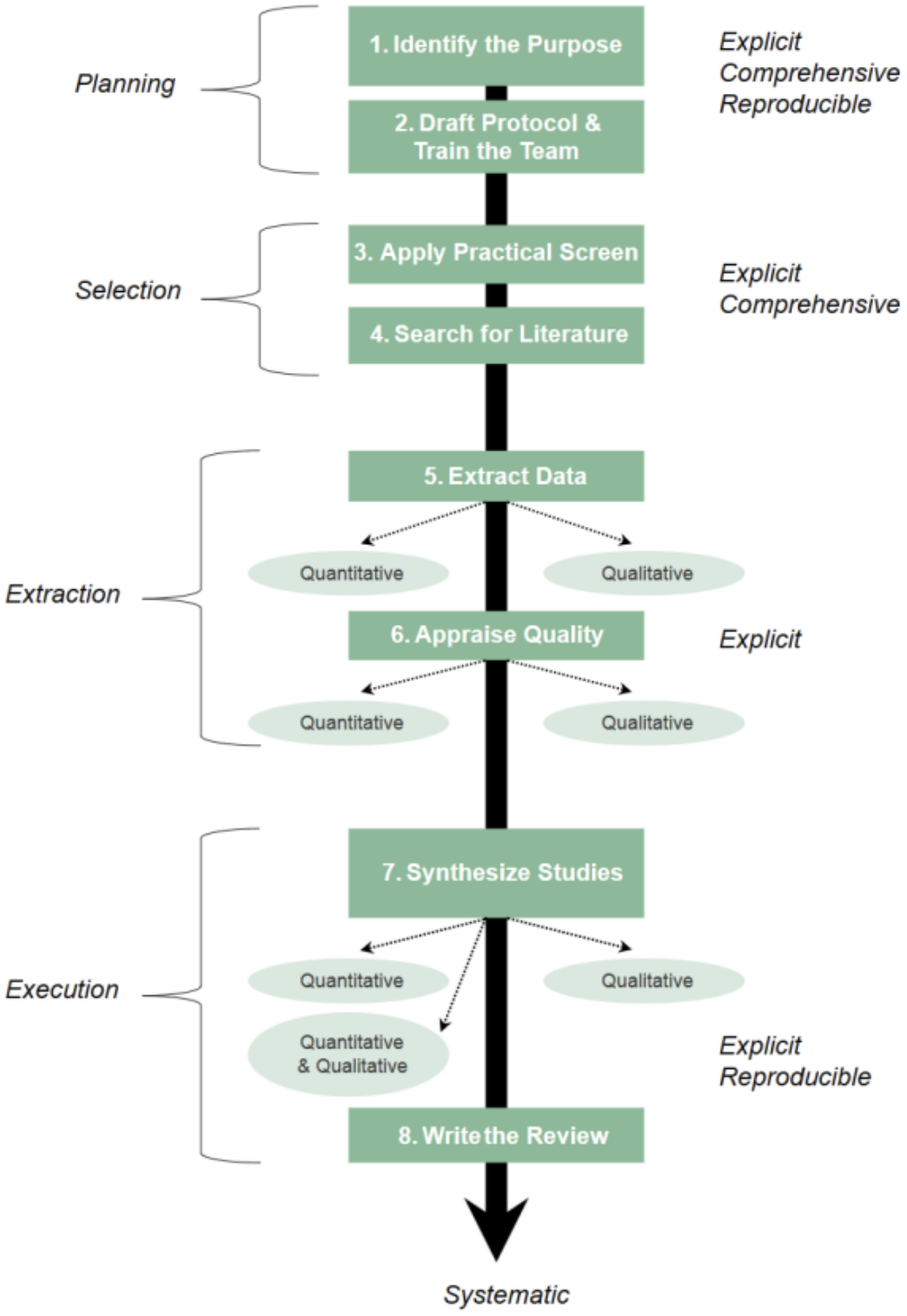
explain at least one solution for each step in the review process and provide an overview of the different approaches for each step in the review. Further, they are planning to use a deep learning approach to aid the process of primary study selection.

Da Silva Júnior & Dutra (2021) propose a general framework for SLR resulting from an overview of existing artificial intelligence tools in four main areas, thus the search-, screening-, extraction-process and synthesis. The authors compare automation approaches and suggest query expansion and term extraction to improve the search process, text clustering and classification and citation mining to improve the screening process, information extraction and visualisation for the extraction process and multi-document summarisation and natural language generation for synthesis. Ultimately, they suggest that the use of supervised deep learning algorithms will potentially lead to higher quality of automation. Nevertheless, they state that many manual processes are still necessary, due to a lack of wholistic tools and that hybrid approaches, by combining rule-based approaches and machine learning are most useful in the current context of SLRs.

Okoli (2015) discuss when a SLR is useful and distinguish between 3 types of reviews. The theoretical background, which provides the theoretical foundation and context of a research question and aids to bring the question into focus. The thesis literature review of a graduate thesis and the standalone literature review, which is a journal length paper in a field without the collection or analysis of primary data. Furthermore, the author describes an eight-step process to guide the SLR (see Fig. 5 and 6). Figure 5 groups the eight steps, differentiates between qualitative and quantitative reviews and provides the order of the process. Figure 6 provides a detailed description of each of the eight steps in the process of SLR. This outline can be utilised as a checklist.

Figure 5

Systematic Guide to Literature Review Development



Note. A Systematic Guide to Literature Review Development adapted from Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1954824

Figure 6

Process of Systematic Literature Reviews

1. **Identify the purpose:** the first step in any review requires the reviewers to clearly identify the review's purpose and intended goals, which is necessary for the review to be explicit to its readers.
2. **Draft protocol and train the team:** for any review that employs more than one reviewer, reviewers need to be completely clear and in agreement about the procedure they will follow, which requires both a written, detailed protocol document and training for all reviewers to ensure consistency in how they execute the review.
3. **Apply practical screen:** also called screening for inclusion, this step requires that the reviewers be explicit about what studies they considered for review and which ones they eliminated without further examination (a very necessary part of any literature review). For excluded studies, the reviewers must state their practical reasons for not considering them and justify how the resulting review can still be comprehensive given the practical exclusion criteria.
4. **Search for literature:** the reviewers need to be explicit in describing the details of the literature search and need to explain and justify how they assured the search's comprehensiveness.
5. **Extract data:** after reviewers have identified all the studies that should be included in the review, they need to systematically extract the applicable information from each study.
6. **Appraise quality:** also called screening for exclusion, the reviewers need to explicitly spell out the criteria they use to judge which papers they will exclude for insufficient quality. Researchers need to score all included papers, depending on the research methodologies they employ, for their quality.
7. **Synthesize studies:** also known as analysis, this step involves combining the facts extracted from the studies by using appropriate techniques, whether quantitative, qualitative, or both.
8. **Write the review:** in addition to the standard principles to be followed in writing research papers, the process of a systematic literature review needs to be reported in sufficient detail such that other researchers can independently reproduce the review's results.

Note. Eight-step process of Systematic Literature Reviews adapted from Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1954824

The most prominent general framework has been developed by Petticrew and Roberts (2006a). Their twelve-step process is the most detailed and rigorous framework that can be used to structure a SLR (see Fig. 7). In their book the authors describe each step of the process in detail and provide suggestions to adhere to high quality standards.

Figure 7

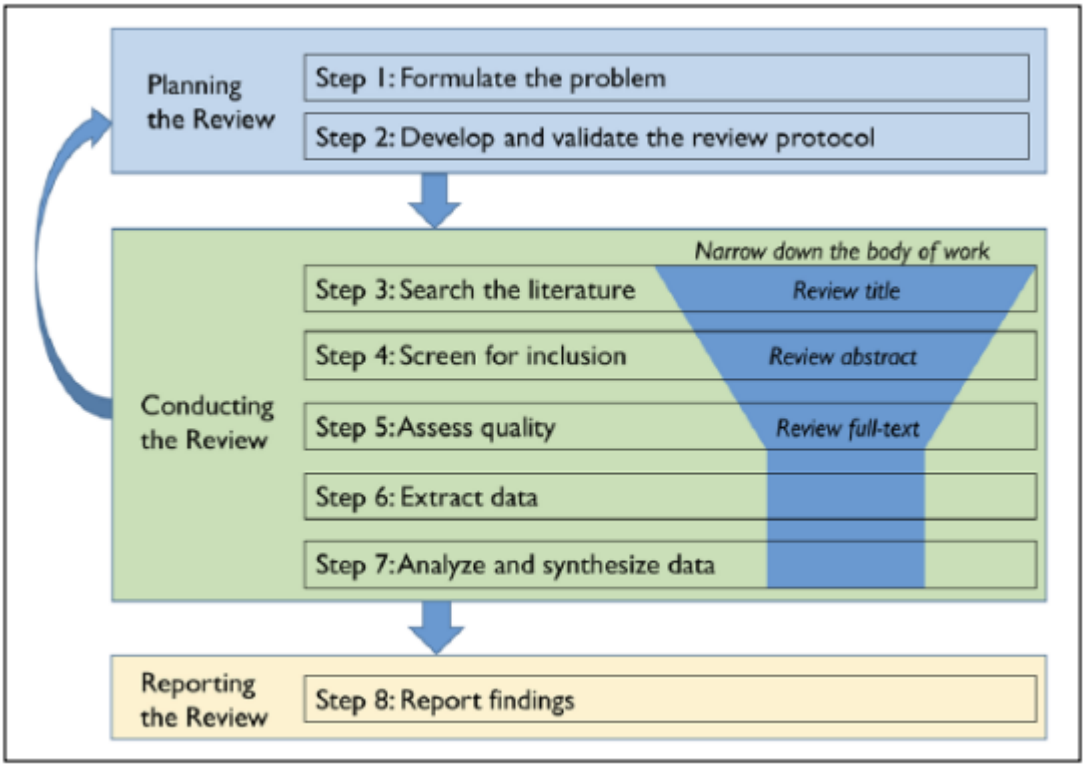
Systematic Literature Review Process of Petticrew and Roberts (2006)

Note. The systematic literature review process developed by Petticrew and Roberts (2006).

The typologies of literature reviews are categorised by Xiao and Watson (2019). Further, the authors discuss steps in conducting a systematic literature review (see, Figure 8). Moreover, the authors provide suggestions on how to enhance rigor in literature reviews for planning in education and research using a funnel to narrow down the body of work.

Figure 8

Process of Systematic Literature Reviews



Note. Process of Systematic Literature Review adapted from Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research*, *39*(1), 93–112. https://doi.org/10.1177/0739456X17723971

The EBSE Technical report (Kitchenham, 2007) provides a detailed overview of steps to take in the SLR process aimed at the domain of software engineering. They developed quality checklists for quantitative and qualitative studies. This approach is commonly used in addition to the use of a general process, which are explicated above.

Guidelines

When automating the procedural aspects of SLR the first step is to standardise steps in the review process. Therefore, guidelines have been developed to organise the processes of SLRs. Various guidelines for SLR have been developed for different disciplines. Bai et al. (2019) conduct a systematic literature review on the evolution of guidelines for SLR. They conclude that the aim of all review guidelines is to support researchers to produce a high-

quality review (Bai et al., 2019). The authors synthesised from literature that most guideline suggestions follow a structured approach. The fundamental structures are determining the focus and scope of the review, developing criteria for inclusion and exclusion, collecting and analysing relevant studies, developing a theoretical framework and reporting on the finding. On the one hand the quality of literature is highly associated with the notion of rigor (Bandara et al., 2015), which is expressed in a highly structured review process, by offering very specific guidelines for conducting certain stages of the review. On the other hand, the aim of the review is to develop and understand insights through continuous engagement with the reviewing process. Wolfswinkel et al. (2013) propose a grounded theory approach, which is a five-step process, consisting of a define-, search-, select-, analyse- and present-stage that enables researchers to come up with theory-based or concept-centric accurate reviews. Ultimately, the review protocol should be replicable, transparent and unbiased, hence it is necessary to record and motivate the procedures. Okoli (2015) developed a guide for a rigorous and standardised methodology for conducting SLR.

The systematic literature review specification manual proposed by Glade (2008) provides a two-step method for conducting a comprehensive review of all types of scientific evidence relating to a question of interest, employing an inclusive, unprejudiced approach. First, all possibly relevant literature is assembled, brought into a standard format, but is not interpreted or any conclusions are drawn upon it. Second, the evidence is assessed by independent experts who did not participate in the collection process followed by an external peer review. The purpose of the manual is to “provide a means of ensuring that data are explored, analysed, and displayed systematically and in a common format to make their assessment and judgments based on this assessment as reliable as possible” (Glade, 2008, p.2).

An ideal article for Webster and Watson (2002), motivates the research topic and explains the contributions of the review. Further, it describes the key concepts, delineates the boundaries of the research and reviews relevant prior literature. Moreover, it develops a model to guide future research, justifies propositions by presenting theoretical explanations, past empirical findings, and practical examples. Ultimately, it presents concluding implications for researchers and managers. In addition to the aforementioned criteria the exemplary review article should be of explanatory and creative nature. The authors developed a step-wise guide on how to conduct SLR.

Domain Specific Issues

The SLR process differs to a degree in each research domain. Therefore, it is important to adapt the SLR process for each research inquiry. For example, Brereton et al. (2006) applied the SLR process in the software engineering domain and identified several issues which need to be addressed in the process such as abstract quality, search terms and database infrastructure. Further the authors address the quality assessment and discuss that keywords are not standardised. Moreover, it is deemed that protocol piloting activities with a formal evaluation are a necessity as well as keeping a detailed record of decision throughout the review process. Furthermore, each member of the research team should be involved in the protocol creation in order to be on the same page throughout the whole process. Cook et al. (1997) on the other hand adapt the SLR process for the health care domain, where SLRs are increasingly used to inform medical decision making, to plan future research agendas, and to establish clinical policy. Their aim is to strengthen the link between best research evidence and optimal health care. Moreover, it is also relevant to take quantitative data into consideration when making clinical decisions. Therefore, replication studies, meta-analyses and integrated datasets are also of relevance in that domain. A general SLR / Meta-Analysis tool supporting all of these features can be deemed necessary.

Current Level of Uptake

Research carried out by van Altena et al. (2018) investigated the current level of uptake of SLR automation tools and what barriers and facilitators exist for the adoption of SLR automation tools in systematic reviews. They identified several tools and asked researchers which tools they were using and why. The results mostly confirm many of the conclusions and recommendations of the ICASR meeting in 2016. The benefits of tools are experienced either through first-hand experience or through experience of colleagues, however, there is a shortage of formal validations of tools. This "points towards a strong influence of the environment and community for improving awareness, evaluation and support for tools" (van Altena et al., 2018, p. 9). Furthermore, a joint development of the validation of tools and the quality criteria need to be developed in order to assess the output of automation tools.

PRISMA Statement

The preferred reporting items for systematic reviews and meta-analyses (PRISMA) (https://www.prisma-statement.org/documents/PRISMA-P-checklist.pdf, retrieved 04.03.2022) statement was designed to aid systematic reviewers to transparently report on the

process and results and recommends the use of the PRISMA flow diagram to depict the flow of studies through the different phases of the systematic review (Kahale et al., 2021). The PRISMA flow diagram is a widely used tool for reporting original systematic reviews. The PRISMA statement is extended for living systematic literature reviews by Kahale et al. (2021).

PICO Framework

The PICO framework was developed to structure clinical questions by capturing the key elements required for a focused research question, which should concern population or problem, intervention, control and outcome.

Support of Management of Systematic Literature Reviews

Many of the difficulties in SLRs are related to administrative complexities that occur when managing and controlling any large and complex project (Bowes et al., 2012). In order to support reviewer in the process and to aid in the reporting of important information, like the number of publications included or the inter-rater reliability, the authors developed a tool called SLuRp. This tool is currently not supported anymore. However, tools such as Covidence have similar features and its uptake is increasing.

Alternative Approaches

Khangura et al. (2012) develop in the context of the Knowledge to Action research program an eight-step approach for producing rapid reviews. The approach addresses a need by knowledge users for timely, user-friendly and trustworthy evidence. The main difference between SLRs and rapid reviews is the time it takes to conduct, which is less than 5 weeks for rapid reviews and 6 to 24 months for SLRs. Nevertheless, the standardised process of rapid reviews is similar to SLRs, they mainly differ in the extend of the search and interpretation of the evidence. Rapid reviews are a way to synthesise evidence to inform evidence-based decision making. Furthermore, they are less costly than SLR which cost are estimated to cost approximately $140.000 (Michelson & Reuter, 2019).

Watson and Webster (2020) suggest that researchers should consider the use of appropriate mindfulness training to enhance their creativity, which could help in synthesising existing knowledge and in proposing future research directions. This is an alternative approach that could be integrated into every workflow.

Search Process

The Search process is concerned with finding relevant research articles to answer the research question. This is commonly carried out by using search queries in several databases. Several methods for improving this process have been suggested in the literature. Using semantic facets (Atanassova & Bertin, 2014), text similarity analysis (Jia & Liu, 2018), development of a checklist (Bin Ali & Usman, 2018; vom Brocke, 2015), using full-text searches (Blanc et al., 2015; Nováček et al., 2009), indexing of research articles by using taggers (Cohen et al., 2021), semantic indexing and similarity queries (Koukal et al., 2014), automation and bibliometrics (Pulsisri & Vatananan-Thesenvitz, 2018), meta-search engine (Sturm et al., 2015), proposing design principles (Sturm & Sunyaev, 2019; Schoormann et al., 2018), defining meta-requirements (Sturm & Sunyaev, 2017), using graph theory (Watson & Webster, 2020), using Application Programming Interfaces (API) and scrapers (Goldfarb-Tarrant et al., 2020) or machine learning techniques (Marshall & Wallace, 2019; Ros et al., 2017) can improve the search process. Furthermore, software solutions exist which can support the search process. These are called connectedpapers.com and elicit.org.

Software Solutions

Next to the established and widely used search engines a software solution called Connected Papers enables researchers to search for specific terms e.g., keyword or specific papers and provides an overview of related papers in the field. It is an improved version of citation networks, which also takes in conceptual and subject area information into account to create a visualisation in a mind map form, which can be read easily.

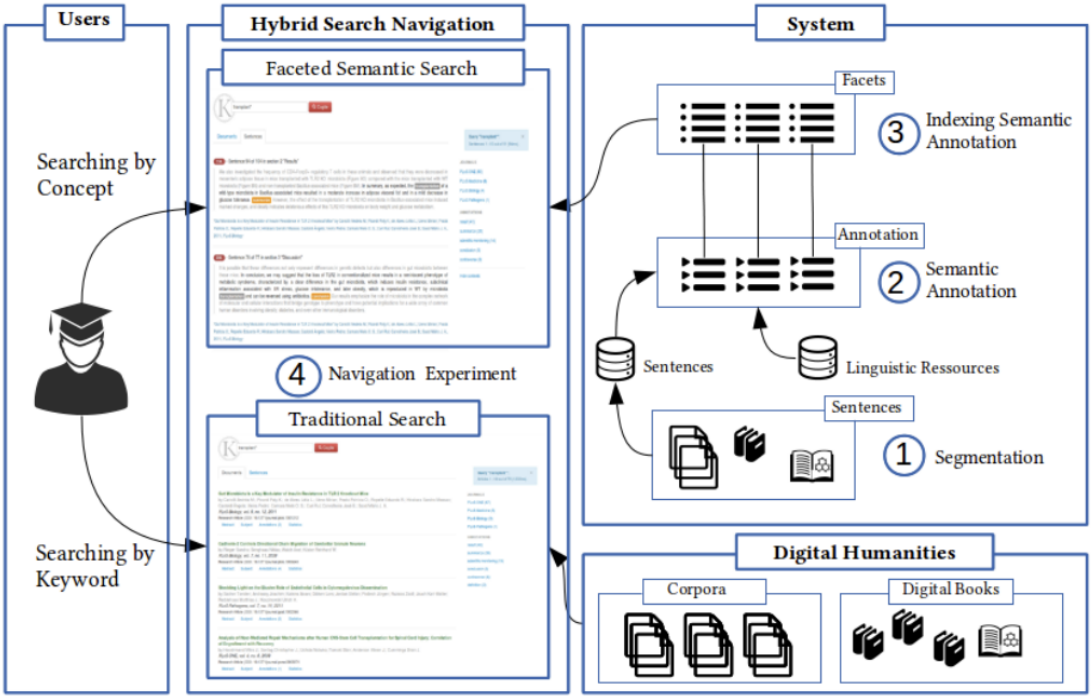
Elicit is a GPT 3-based search engine for research papers, which enables researchers to ask questions to the research paper, whereby the answers to these questions are provided in a table format. Asking questions, thus conducting queries about information to a research paper enables enormous opportunities in generating meta-data about research papers in order to gain an overview of multiple research articles before even reading a paper. Nevertheless, the tool is still in development and the error rate is still too high to be feasible for basing research results on it. It still requires a lot of cleaning from researchers. However, the quality of the output of the tool is improving rapidly. The search results can be downloaded in a .bib file, which can be imported into a reference manager software like Zotero. Levy and Ellis (2006) provide guidelines on organising electronic- and paper-based reference management systems.

Semantic Approaches

Atanassova & Bertin (2014) filter search results of scientific papers according to semantic facets. The semantic annotations are obtained using a rule-based method that identifies specific linguistic clues organized into a linguistic ontology. Paragraphs in the dataset were segmented into sentences. It offers efficient search and navigation in scientific papers. The full overview of the study can be found in Figure 9 of Atanassova and Bertin (2014).

Figure 9

Faceting and Semantic Annotation



Note. This figure illustrates faceting and semantic annotation of research publications adapted from Atanassova, I., & Bertin, M. (2014). Faceted Semantic Search for Scientific Papers. In *Semantic Publishing Challenge co-located with the 11th European Semantic Web Conference ESWC2014*.

The author's goal is to develop other semantic categories and facets related to scientific articles. Further improvements in the segmentation and annotation process of text can help to support the search process even further. Tsafnat et al. (2014) suggest synonym

expansion and word sense disambiguation to improve the search process. Feng et al. (2017) discuss the federated search approach to provide a unified interface for all defined data sources and the query expansion technique. Jia & Liu (2018) on the other hand explore the expansion of search strings using text similarity analysis. First, the basic search string is derived from the research question and segmented into words. Second, for each notional word (its synonyms and hypernyms) are retrieved from WordNet to obtain new keywords. A combination of these new keywords creates a new search string, by calculating the similarity between the new strings and the basic string. Suitable new search strings with high similarity are then selected and used in the search process. If a paper is selected it is converted into word vectors and the similarity between the words and vectors and the research question is calculated.

Koukal et al. (2014) proposed a Tool for Semantic Indexing and Similarity Queries (TSISQ) which is built upon the well-founded technique called latent semantic indexing (LSI). LSI, also called latent semantic analysis (LSA) is part of the natural language processing techniques, which maps meaning into a semantic space. Another theory that the TSISQ builds upon is the query expansion (QE), which is an information retrieval technique that aims to advance retrieval effectiveness. This is done by extending the provided search terms with synonyms or related terms (Koukal et al., 2014). The TSISQ prototype "enables researchers to efficiently gain an overview of a specific research field, deepen their knowledge and furthermore, to refine the theoretical foundations of their research" (Koukal et al., 2014, p.8). The authors concluded that the approach can save time in finding basic literature and to help to increase the comprehensiveness of the review by identifying sources that might not have been taken into account using the classic methodology.

Text Similarity Analysis

The expansion of search terms via text similarity analysis can be applied in two stages of the SLR (Jia & Liu, 2018). The authors propose the use of synonym and hypernym relations based on a universal dictionary such as WordNet to expand the search strings, which can improve the search process. The authors further suggest to calculate the relevance of research articles according to text similarity analysis to improve the primary study selection.

Checklists and Design Principles

Bin Ali and Usman (2018) develop a comprehensive evaluation checklist to assess the reliability of an automated search strategy used in an SLR. The two dimensions of reliability

are repeatability and consistency. Repeatability means sufficient details are reported to redo the search and identify the same papers. Consistency means how rigorous the search has been conducted. Both rely on the transparent and auditable reporting of the search process (Bin Ali & Usman, 2018).

Challenges of literature searches in the increasingly dynamic context are discussed by vom Brocke et al. (2015). The authors present detailed practical guidelines and a checklist to help researchers with planning and organizing their literature searches. Vom Brocke et al. (2015) discuss the iterative search process, by going back and forth from searching and reading literature. Further, the authors suggest that the narrative review style is more suited for experienced scholars, while less experienced researchers are likely to benefit from more a systematic and guided approach. Ultimately, the author state that here is no one-size fits all approach in search, because it can be difficult to evaluate the quality and relevance of publications, the retrieval and storage of literature is mostly messy and time consuming, it is very difficult to tell when the search is finished. Therefore, vom Brocke et al. (2015) suggest to develop an understanding of the subject matter before starting the search process, to justify the review and the literature search, to define the scope of the search, to test and apply search parameters, to use seminal sources to build the structure of the review and to weigh the feasibility of the search against their coverage. These suggestions can be summarised by stating that it is necessary to develop a protocol to guide and document the search process, therefore the authors have developed a literature search checklist which is presented in Table 5.

Table 5

Literature Search Checklist

Before the literature search	Checked
Develop an understanding of the topic	
Justify the necessity of the literature review	
Define an appropriate research scope	
Assess the feasibility and coverage of the search	
During the literature search	
Test alternative approaches to searching literature	
Use justifiable search techniques and parameters	
Apply appropriate criteria for inclusion and exclusion	

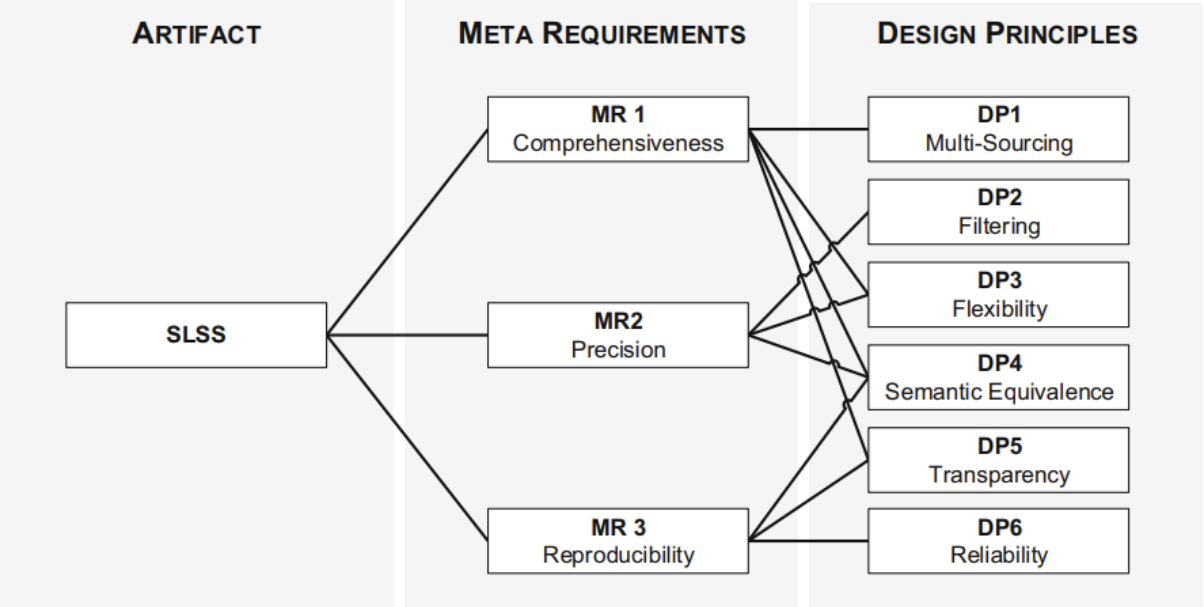
Discuss the search strategies in the team
After the literature search
Assess the sensitivity and specificity of the search
Rigorously document the search process and results
Compare the results to those of other literature reviews
Collect feedback from colleagues

Note. This table is adapted from vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the Association for Information Systems*, *37*. https://doi.org/10.17705/1CAIS.03709

Sturm and Sunyaev (2019) describe six design principles to summarise the understanding of effective SLR research design, that are derived from multiple design cycles that intend to increase the comprehensiveness, precision, and reproducibility of systematic literature searches. The design principles are multi-sourcing, filtering, flexibility, semantic equivalence, transparency and reliability. The meta-requirements are comprehensiveness, precision and reproducibility of the search process (Sturm & Sunyaev, 2017). Figure 10 in Sturm and Sunyaev (2018) provides a visual overview.

Figure 10

Mapping of Systematic Literature Search Systems Meta-requirements and Design Principles



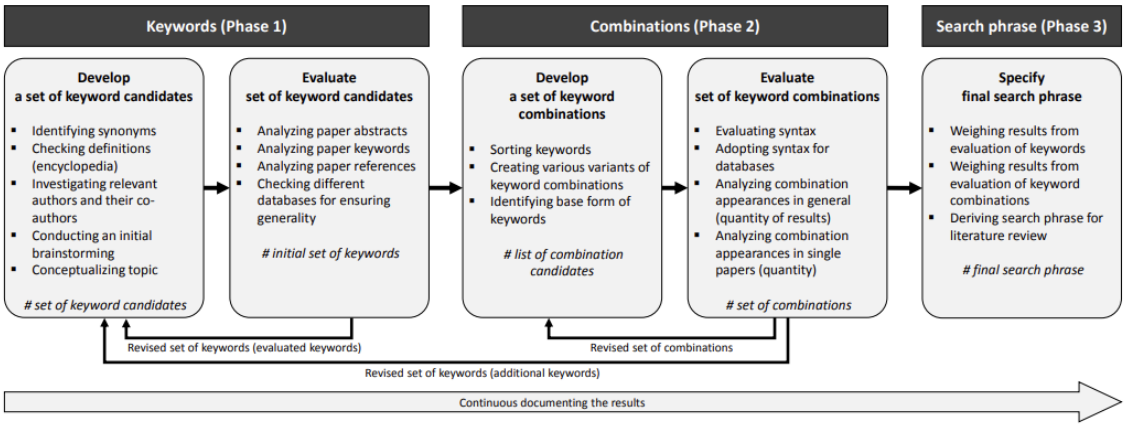
Note. This figure illustrates the mapping of SLSS meta-requirements on SLSS design principles adapted from Sturm, B., & Sunyaev, A. (2019). Design principles for systematic search systems: A holistic synthesis of a rigorous multi-cycle design science research journey. *Business & Information Systems Engineering, 61(1),* 91–111. https://doi.org/10.1007/s12599-018-0569-6

The design principles were instantiated in form of an application called LitSonar, which is able to conduct meta-searches on several databases. Sturm et al. (2015) developed LitSonar for academic literature which consolidates search results from several literature databases. Developed with an incremental development approach consisting of multiple design cycles of artefact creation / refinement and qualitative / quantitative evaluation. It enables the search of keywords in several databases at the same time using publication and ranking filters.

Schoormann et al. (2018) propose a set of design principles for the search process and translate them into key features that can be implemented in the form of concrete IT-artifacts. First of all, it is very important to have a process set up to document the search. Second, guidance though the search process steps for the derivation of a search phrase should be provided, the model for specifying the search phrase can be found in Figure 11 of Schoormann et al. (2018).

Figure 11

Derived Procedure Model for Specifying a Search Phase



Note. This figure illustrates the derived procedure model for specifying a search phrase adapted from Schoormann, T., Behrens, D., Fellmann, M., & Knackstedt, R. (2018). Design principles for supporting rigorous search strategies in literature reviews. *2018 IEEE 20th Conference on Business Informatics (CBI)*, 99–108. https://doi.org/10.1109/CBI.2018.00020

Third, a support tool that automatically generates various search phrases based on the syntax rules of a database, that a researcher can select from. Fourth, a tool that tracks the entire search and generates a written document which can be included in the report. Ultimately, the documentation and extraction allow more rigor in the description and retraction of the search process.

Zhang et al. (2011) aim to validate the QGS-based (Quasi-gold-standard) search process for SLR. The authors developed a five-step process. The first is to identify related venues and libraries, the second is to establish the gold standard by screening all articles, the third is to elicit search strings for query, the fourth is to conduct automated search using specific syntax and criteria of each library and the fifth is to evaluate search performance using the precision and sensitivity metrics. Therefore, it is an empirical assessment of the systematic search process. The authors further postulate that an integrated search strategy will retrieve more relevant studies than a manual, automated or combined search strategy for a SLR. The process uses the results from one search method to inform the design or evaluation of another search method.

Indexing Mechanisms

Blanc et al. (2015) assess the efficiency and reliability of full-text searches in major medical journals for identifying shared decision-making publications. Their major finding was that indexing mechanisms for full-text search is an essential component to sort text. This is a task requiring a lot of manual-task completion. Nevertheless, (Cohen et al., 2021) propose a multi-tagger approach for automatically indexing articles according to publication types (PTs). Study design taggers can be an aid to filtering literature for information retrieval purposes. Multiple features were included in the training of the automated machine learning based probabilistic PT and study design tagger model, which are title, abstract and other meta-data. It can support the evidence synthesis process by grouping evidences from literature. The multi-tagger approach could potentially assist in automated ranking of articles for inclusion. (Ros et al., 2017) Propose a machine learning approach to support semi-automated selection in SLR by using classifiers to identify research articles.

Nováček et al. (2009) conduct full-text queries and knowledge-based queries in CORAAL (COntent extended by emeRgent and Asserted Annotations of Linked publication data). It extracts asserted publication meta-data together with the knowledge implicitly present in the respective text, integrates the emergent content and exposes it via a multiple-perspective search & browse interface.

Goldfarb-Tarrant et al. (2020) discuss three main stages of a literature review that can be carried out automatically. These are the searching for documents than can be done via Application Programming Interfaces (API) and scrapers and selection of relevant documents using binary classification and extraction of data via sequence-labelling classification based on human-expert annotation. Therefore, they are classifying documents and carry out the data extraction to provide accurate search results which can easily be selected from.

Bibliometrics

Pulsisri and Vatananan-Thesenvitz (2018) improve the search and screening process by using automation and bibliometrics. The authors describe automation as the operationalisation of the manual tasks of the SLR using specific tools and computer systems. Bibliometrics are described as the process of analysing the bibliographic data of published literature to provide an overview of the body of knowledge for a given field of inquiry. They discuss the CIMO model (discuss, intervention, mechanism and outcome) which can be applied for research question formulation. Marshall & Wallace (2019) provide an overview

of current machine learning methods to improve the search, selection and data extraction processes and offer guidance on how to use them in practice. Text classification and data extraction are the core of natural language processing technologies for SLRs. The authors highlight the fact that the use of technology is interwoven with manual tasks and conclude that it is very unlikely, in the short term, to have full range automation tools. Therefore, human-machine interaction are key features of current SLRs.

Wells (2016) discusses discovery systems in the domain of libraries. The author describes the four user tasks and core functions, which are, find, identify, select, obtain and address the interaction with social media and other websites as a way to enrich discovery systems.

Graph Theory

Watson and Webster (2020) have been discussing methods of improving the search and selection process by using graph theory, by formally mapping the relationship among core elements, like concepts and processes of a synthesis, with the aim of gradually building a meta-synthesis of Information Systems research.

Comparison of Search Systems

Gusenbauer and Haddaway (2019) evaluated which academic search systems are suitable for systematic reviews or meta-analyses. The authors evaluated retrieval qualities of Google Scholar, PubMed, and 26 other resources. The authors distinguished between academic search systems regarding coverage, search strings and queries, search results and search reproducibility, necessary and desired criteria. Their future outlook discusses the emergence of semantic search engines and their promising character.

Screening Process

The screening process is concerned with identifying relevant papers to answer the research question. Three sources for decision-making are available, citation screening, abstract screening, full-text screening, figure screening. The manual systematic literature review screening process follows a sequence of steps. First the potentially relevant papers' title resulting from the search process are determined to be relevant or not by at least two researchers. Second, the determined relevant papers abstracts are examined and determined to be relevant or not by at least two researchers. Third, the determined relevant papers' full text is examined in preparation for the selection and data extraction process. Several tools can be used to do this collaborative work like Covidence and more. Several approaches to automate

the screening step in the systematic literature review process have been identified. Machine learning approach for automating text classification approach (Bekhuis & Demner-Fushman, 2010; Jaspers et al., 2018), automated text annotation approaches (Desclés, 2006), Text mining (Feng et al., 2017; Rathbone et al., 2015), evaluation approach by setting benchmarks (Lange & Di lorio, 2014), crowdsourcing for citation or abstract screening (Mortensen et al., 2017; Lee et al., 2017; Nama et al., 2019), provide an overview of the body of knowledge (Pulsisri & Vatananan-Thesenvitz, 2018), use of summaries of research articles (Uban & Caragea, 2021; Ibrahim Altmami and El Bachir Menai, 2020; Erera et al., 2019).

Machine Learning Approaches

Bekhuis & Demner-Fushman (2010) utilise the supervised machine learning method EvoSVM to identify relevant papers, by trying to classify research papers according to their study design. The authors conclude that supervised machine learning can reduce the workload during the initial screening phase, because the tool reduces the amount of potentially relevant studies for the first reviewer of the literature.

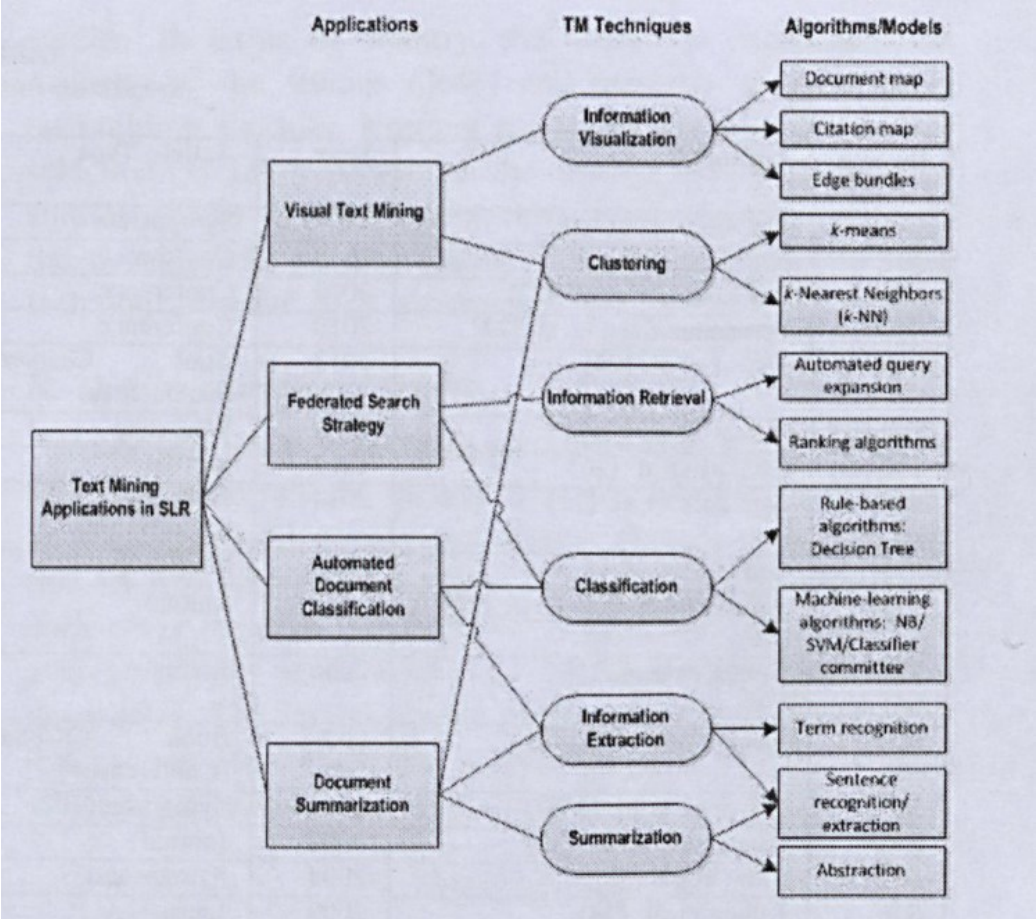
Jaspers et al. (2018) utilise machine learning techniques for the screening process. The authors developed a shiny application with R that focusses on abstract and full text screening, that can be used in combination with other tools such as DistillerSR. Classifiers are being trained by using a labelled dataset created by human reviewers, which are then used to label other research papers. This process is the same for full text analysis, however different classifiers have to be used.

Text-Mining Approaches

Two approaches for text mining are mentioned by Feng et al. (2017). First, the automatic term-recognition approach and second, the classifier construction, which can either be rule-based or active ML techniques. Figure 12 developed by Feng et al. (2017) shows the main text-mining applications in SLR. The authors outline that citation and full-text screening can be conducted using text mining by clustering, classification and summarising the information in a text to sort research articles by relevance.

Figure 12

Main Text-Mining Applications in Systematic Literature Reviews



Note. This figure illustrates the main text-mining applications in Systematic Literature Reviews adapted from Feng, L., Chiam, Y. K., & Lo, S. K. (2017). Text-mining techniques and tools for systematic literature reviews: A systematic literature review. 2017 *24th Asia-Pacific Software Engineering Conference (APSEC),* 41–50. https://doi.org/10.1109/APSEC.2017.10

Rathbone et al. (2015) developed a tool called Abstrackr to semi-automated title and abstract screening of research articles using text mining techniques, by identifying relevant studies. They use of 4 systematic reviews as a basis to evaluate Abstrackr. They found an improvement of efficiency in terms of time saved between 9 and 80% in the different reviews. Rathbone et al. (2015) conclude that it is not yet possible to use it as a stand-alone tool. Nevertheless, it could support the review process of the second reviewer.

Automatic text classification (ATC) can be used in the search and selection process. Desclés (2006) developed a method of annotation called Contextual Exploration which takes

into account context. The research article describes the process of automatic annotation of texts. The annotation happens as an action if a condition has been met. A semantic map emerges that shows the relation between concepts and linguistic units are defined. The map shows the unary concept with binary relations, with different functional types and arrows expressing specifications or generalisations. Expressions are linguistic indices that yield sufficient information for giving a label to a contextual unit. The platform EXCOM has been developed to automate the annotation of texts.

Crowdsourcing

Crowdsourcing can be an effective approach to citation screening for SLRs (Mortensen et al., 2017; Lee et al., 2017; Nama et al., 2019). The authors use crowds recruited via the internet to label and annotate papers for citation screening, creating a hybrid approach using novice recruitees and rule-based decisions. This process is far more cost efficient, a reduction to one sixth of the cost, in comparison to experts doing the same process (Mortensen et al., 2017).

Summaries

Another approach is the use of summaries of research articles to aid in the identification of relevant articles, which can be used on top of the title and abstract. Uban and Caragea (2021) explore automatic review summary generation for scientific papers and evaluate the state-of-the-art neural summarisation models. Text generation models for summarisation are generally neural and transformer based. The fine-tuned extraction summarisation model performed better in full text summarisation than the pre-trained model. The authors suggest to utilise multi-tasking learning to incorporate abstract and full text analysis and conditional text generation as approaches worth researching (Uban & Caragea, 2021). Ibrahim Altmami and El Bachir Menai (2020) did a survey about summarizing scientific articles. The authors conclude that a summary is expected to be informative enough to cover all the main sections of the input article and that it should present the most important information the reader is looking for. Erera et al. (2019) developed a system called IBM Science Summarizer, which creates summaries of scientific articles.

Pulsisri and Vatananan-Thesenvitz (2018) improve the screening process by operationalising the manual tasks of the SLR using specific tools, computer systems and bibliometrics to provide an overview of the body of knowledge for a given field of inquiry.

The natural language model GPT-3 is able to output simple summaries of text. Although many limitations apply to this approach, it is a great first step to support researchers in their work. The author of this thesis played around with it and tested out some of the possible applications and tested for feasibility. However, the model was unfortunately not able to help writing this thesis using simple commands. Nevertheless, it is very promising and worth it to consider spending more time on the model to test out its limits and how to push the right commands for the tasks of SLR. The tool Elicit is based on the GPT-3 model and is able to do some text extraction tasks among many others.

Software Solutions

Covidence is an open-sourced web-based software platform that streamlines the production of systematic reviews, including Cochrane Reviews. The solution offers the features citation screening, full text review, risk of bias assessment, extraction of study characteristics and other study data and is able to export data into citation management platforms, like RevMan.

Selection Process

The selection process is concerned with choosing which research papers to include in the literature review and data extraction section. Several approaches have been identified such as visual text mining techniques (Felizardo et al., 2012), concept maps (Dos Santos et al., 2017), development of a protocol for article selection (Ferreira et al., 2021), node-link graph (Chou & Yang, 2011), citation network approach and automated detection of implicit theory (Larsen et al., 2019), crowdsourcing approaches (Nama et al., 2019; Lee et al., 2017; Weiss, 2016), development of guidelines (Nama et al., 2019), semi-automated machine learning approach (Ros et al., 2017) a linked data approach (Tomassetti et al., 2020) and quality evaluation criteria (Lange & Di lorio, 2014; Leidner, 2018).

Visual Text Mining

Felizardo et al. (2012) discuss the use of visual text mining (VTM) techniques for primary study selection. They propose a novel approach to support the primary study selection activity using VTM techniques. They created a document map, edge bundles, clusters and topics, expression occurrence and a citation network. The pilot study showed VTM reduced the time spend on the search process and increased the number of studies correctly included. Furthermore, it is necessary for SLRs to be updated. Therefore, Felizardo et al. (2014) developed a visual analysis approach called USR-VTM, which builds on

authors' previous work. This allows to generate a data set that can be updated, which in turn creates the ability to update SLRs.

Concept Maps

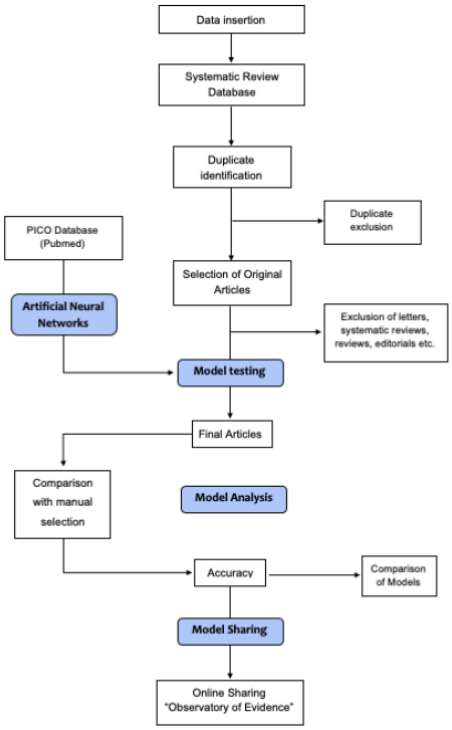
Dos Santos et al. (2017) discuss the use of Concept Maps (CM) for identifying relevant studies, summarise a complex structure of textual information. CMs have a flexible structure, are easy to understand and allow for knowledge sharing. However, CMs are limited in their ability to scale to an arbitrary size. Many examples of CM are provided by Petticrew and Roberts (2006).

Protocol for Article Selection

Ferreira et al. (2021) developed a protocol for an article selection model. They created models based on an artificial neural network system to automate the article selection process in systematic reviews. The following steps have been taken: data import, exclusion of duplicates, exclusion of non-articles, article reading and model creation using artificial neural network, comparison of the models and system sharing, which can be seen in more detail in Figure 13 of their research article.

Figure 13

Flowchart of the Automated Steps in Article Selection System for Systematic Reviews



Note. This figure illustrates the flowchart of the automated steps in article selection system for systematic reviews generated from the PICO statement: Population, Intervention, Control and Outcome adapted from Ferreira, G. F., Quiles, M. G., Nazaré, T. S., Rezende, S. O., & Demarzo, M. (2021). Automation of article selection process in systematic reviews through artificial neural network modeling and machine learning: Protocol for an article selection model. *JMIR Research Protocols*, *10*(6), e26448. https://doi.org/10.2196/26448

Graphical Representation

PaperVis is a tool developed by Chou and Yang (2011) that arranges papers in a node-link graph to depict the relationships between papers. Expert users can find important papers in the specified categories by using keywords to filter information from the dataset, whereas novice users can discover papers of higher importance in a new research field (Chou & Yang, 2011). The tool represents interesting papers as a graph that depicts their complex relationships by utilising a modified version of existing radial space filling and bullseye view techniques (Chou & Yang, 2011). Furthermore, visual cues, such as the colours, sizes and boundaries of nodes to indicate a papers' importance and relationship with other papers. The distance between papers determines their relative similarity. Lastly, the authors have developed a clustering algorithm to meaningfully categorise and group papers (Chou & Yang, 2011).

Citation Network

Larsen et al. (2019) discussed a citation-network approach to boundary classification and introduced the automated detection of implicit theory (ADIT) technique. ADIT is a design instance which is suggested by the discourse approach. It uses machine learning to select the empirical theory-contributing manuscripts within a theory ecosystem. It has three general steps. The construction of a theory ecosystem to provide a comprehensive set of manuscripts for boundary identification, a random sample is coded and a corpus construction is done by the selection of manuscripts. ADIT improves performance over the conventional approach as practiced in past technology acceptance model reviews (Larsen et al., 2019). The identification process determines the size and delineation of the corpus and the corpus construction process is based on machine learning to classify manuscripts as more or less likely to be relevant.

Crowdsourcing

Crowdsourcing is the practice of obtaining participants to carry out specific tasks, by providing services, ideas and content most commonly via the internet. It allows a large group of people to work on a common goal by part-taking in a process, for example a review process. Innovation tournaments, prizes for solving an engineering problem or paying online participants for categorizing images or any other form of data are examples of crowdsourcing. “Crowdsourcing is an approach to accomplishing a task by opening up its completion to a broad section of the public” (Ranard et al., 2014 p. 1). Several platforms can be used to outsource tasks, such as mTurk which was successfully used by Mortensen et al. (2017) for abstract screening, the Cochrane Crowd initiative and Crowdflower (Nama et al., 2019). Outsourced tasks need to have requirements and standards on who is able to participate in the carrying out certain tasks. Lee et al. (2017) use crowdsourcing for the selection of research articles by citation and abstract screening in the health-care sector.

Nama et al. (2019) utilise crowdsourcing for citation, abstract and full-text screening. They familiarize the crowd with both the SR eligibility criteria and platform. The study flow diagram can be found in Figure X. (Nama et al., 2019, p. 5). They achieved great results utilising crowdsourcing. “Although crowdsourcing has the potential to lead to more rapid knowledge synthesis and evidence translation, it is important to acknowledge that it can only do so if accessible, cost-effective, and scalable” (Nama et al., 2019, p. 12). A few guidelines to achieve good performance are mentioned by the authors. They outline that is important to have a platform that allows researchers to easily access a large and expanding crowd, well-defined eligibility criteria and instructions, quality control measures, performance measures and a performance-oriented incentive structure (Nama et al., 2019). Weiss (2016) utilises crowdsourcing to bring insights into the relevancy of articles for a specific topic, by creating consensus between experts through commenting, voting and tagging.

Classification

Ros et al. (2017) trained a classifier on an initial set of papers, extend this set of papers by automated search and snowballing, have the researcher validate the top paper, which was selected by the classifier. It updates the set of papers and iterate the process until a stopping criterion is met. They implemented this machine learning approach to support semi-automated search and selection of SLRs. It is useful for mapping studies, can be applied for systematic reviews and can be used to automate the updating of included papers.

Tomassetti et al. (2020) implemented a linked data approach to support semi-automated search and selection of SLRs to reduce the workload needed to classify sources and to reduce the subjectivity in the overall process by using existing technologies in the field of Semantic Web and text mining techniques. This approach was able to capture not just primary studies recognize being similar to the ones already selected but was able to capture papers that have conceptual relations to the content expressed in the prior selected papers.

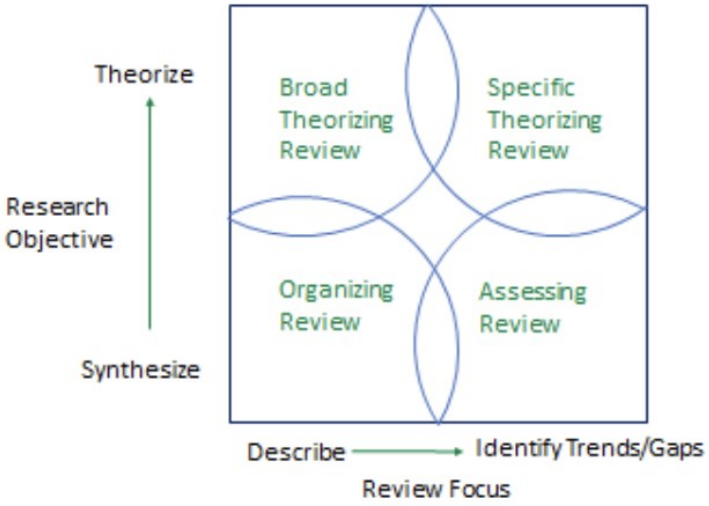
Quality Evaluation Criteria

Lange and Di lorio (2014) created a semantic publishing challenge to evaluate the approaches that other researchers offer. The authors aim to develop an end-user service using Linked Open Datasets (LOS) to identify relevant research papers using information processing techniques, to extract and characterise citations and to assess the quality of scientific output. Several approaches have been evaluated for the three tasks and successful approaches have been developed based on the benchmarks set by the authors.

Leidner (2018) proposes criteria to evaluate the quality of Review and Theory Development (RTD) papers. The author presents a RTD framework to improve the quality of RTD papers. The framework suggests four types of RTD papers. Figure 14 illustrates the characteristics of the four types of review papers with the research objective and the review focus as dimensions. These are organising reviews, assessing reviews, specific-theorising reviews and broad theorising reviews.

Figure 14

Polylithic Framework of Review and Theory Development



Note. A Polylithic Framework of RTD Papers adapted from Leidner, D. (2018). Review and theory symbiosis: An introspective retrospective. *Journal of the Association for Information Systems*, *19*(06), 552–567. https://doi.org/10.17705/1jais.00501

The author concludes that the quality of the review paper depends more on its attributes than on the its type. Ultimately, Leidner (2018) outlines that the challenge is to provide something different and new, yet not so different and new as to render the past unrecognisably, but new and different enough to render the future imaginable.

Software Solutions

Covidence among many other tools, has a collaborative feature that allows to include and exclude identified papers in the search process with a team of researchers, resulting in a list of included papers which can be used as primary sources for SLRs.

Data Extraction Process

The data extraction process is concerned with the selection of relevant information from research articles to answer the research question. Several approaches have been explored, knowledge graphs (Agrawal, 2021), single-document summarisation (Ibrahim Altmami & El Bachir Menai, 2020; Uban & Caragea, 2021), visualisations in an information retrieval context through contextual and cognitive analysis (Bertin & Atanassova, 2012), deep learning approach for classification of references within a full text scholarly publication (Fernandes Rodrigues Alves et al., 2018), semantic annotation (Bertin, 2014), guidelines for organising and preparing papers for analysis and frameworks (Bandara et al., 2015; Jonnalagadda et al., 2015), scientific discourse annotation (Liakata et al., 2012), use of grounded theory (Wolfswinkel et al., 2013), document classification with machine learning classifiers (Hamad & Saiim, 2014), living systematic reviews (Slaughter et al., 2015), identification of research objects (Marshall & Wallace, 2019), software packages such as NVivo.

Identification of Research Objects

Data extraction tools are designed to assist the manual process of SLRs, for example by drawing the user's attention to relevant text or by making suggestions to the user that they may validate (Marshall & Wallace, 2019). Automatic data extraction for systematic reviews means on the one hand, that relevant data from a scholarly publication are extracted to answer the research question, which necessitates the automatic identification of research objects,

hence annotating or labelling of multi-media information. On the other hand, the summarisation of a single document, by minimising the loss of information, but with the highest form of compression. Data extraction tools aim to be utilised as the primary source for data element extraction that would then be validated by a human and could eventually become completely automated to enable living systematic reviews, continuously updated systematic reviews with the latest knowledge available, incorporating the latest publications. Slaughter et al. (2015) discuss necessary next steps towards developing a "living systematic review".

Guidelines

Jonnalagadda et al. (2015) express an "urgent need for a unified framework or system to extract all necessary data elements. Studies need to be conducted using the same gold standard and on the extraction of the same data elements for effective comparison." (Jonnalagadda et al., 2015 p.13). They outline the need for the development of new tools for reporting on and searching for structured data from published literature (Jonnalagadda et al., 2015).

Bandara et al. (2015) introduced a four-phased software tool to conduct literature reviews, by viewing the process as a qualitative study to extract relevant literature and justify its scope, relevance and quality. They describe "how to organize and prepare papers for analysis and provide detailed guidelines for actually coding and analysing papers, including detailed illustrative strategies to effectively write up and present the results" (Bandara et al., 2015, p.1).

Knowledge Graphs

Agrawal (2021) develop a Semi-Supervised Extraction of Structured Information from Scientific Literature by creating structured repositories such as knowledge graphs. Their proposed algorithm was able to extract phrases in a semi-supervised manner from a large dataset of scientific articles. The authors built a graphical representation by using the extracted concepts as entity nodes and the paper nodes that include the metadata of the paper. This is combined with a citation graph representing the relationships between papers. On top of that, relations between entity nodes and paper nodes regarding the aim, method or result are represented. The graph is able to summarise the research community by finding entity nodes corresponding to specific topics, aims or methods. It is also able to summarise the findings of methods used in a specific field or can uncover applications by finding the aims

where the field has been used as a method. Furthermore, it can be used for trend analysis, finding the works of one specific author or to bundle conference proceedings.

Graphical Representation

Felizardo et al. (2011) investigate whether graph representations result in better comprehensibility than tables when researchers are presented with SLR results and to investigate whether the performance is impacted by the interpretation of results using graphs. Three main findings are outlined. First, the graphical representation of SLR data led to a reduction in the time taken for its analysis, without any loss in data comprehensibility. Second, graphical data proved to be faster than the analysis of tabular data. Third, there was not a difference in comprehensibility using the tabular format, the graphical format or a combination. Felizardo et al. (2011) argue that graphs are a suitable alternative to tables when representing results of a SLR.

Summarisation

A single-document summarisation can assist researchers in synthesising multiple documents and multi-document summarisation could help to automate the synthesising aspect of SLR (Ibrahim Altmami & El Bachir Menai, 2020; Uban & Caragea, 2021). However, multi-document summarisation is a complex task and still needs a lot of improvements with respect to understanding and sense-making abilities of the algorithms. Nevertheless, single-document summarisation could be viable addition to abstract screening and can thus support the data extraction process.

Indexing and Annotation

Liakata et al. (2012) discuss three schemes for scientific discourse annotation. The authors identify core components of scientific investigations by focussing on the characterisation of methods, outcomes and objectives, called the SAPIENTA tool, on the characterisation of background work and appropriate segments with the discourse annotation scheme and on the analysis of certain segment and sentence types, which considers the conveyance of epistemic knowledge at the clause level.

Hamad & Saiim (2014) synthesised the following machine learning classifiers for document classification from literature: complement Naïve Bayes, discriminative multinomial naïve Bayes, alternating decision tree, AdaBoost (Logistic Regression), AdaBoost (j48), support vector machine learning algorithm and voting perceptron-based. Moreover, a graph representation can be used to support the data extraction, as well as a meta

search process and text mining to improve the search strategy by using an associative search and a sentence extraction method for multi documents summarization can be used to support the data synthesis process.

Automatic extraction of Information of research papers and semantic analysis provides the opportunity for new types of visualizations in an information retrieval context through contextual and cognitive analysis (Bertin & Atanassova, 2012). The authors propose a linguistic ontology of bibliographic citations in combination with semantic annotation of text and references. If semantic annotation of citations is considered in relation to the document metadata, then better information retrieval and access to the document content is possible. Rodrigues Alves et al. (2018) developed a deep learning approach for the detection, extraction and classification of references within a full text scholarly publication.

Text mining techniques can be used to extract and link named entities from scientific papers, by utilising semantic annotation of corpora using linguistic resources and the citation context (Bertin, 2014).

Grounded Theory

Wolfswinkel et al. (2013) highlight the importance of grounded theory in the SLR process. Grounded Theory "enables the key concepts to surface, instead of being deductively derived beforehand; they emerge during the analytical process of substantive inquiry" (Wolfswinkel et al. 2013, p.46). Grounded Theory forces the reviewers to focus on the researched concepts and is usually based on empirical facts. This is to be done through the fairly random yet methodical ways of "reading the content of all the review's single studies, so that eventually new ideas emerge." For example, through axial coding procedures the review process becomes more of a skill than a subjective art. A good review must be a "richly competent coverage of a well-carved out niche in the literature" (Wolfswinkel et al. 2013, p.47).

Software Solutions

NVivo is a software package produced by QSR International for qualitative analysis. It helps to organise, analyse and find insights in unstructured or qualitative data, where deep levels of analysis of large volumes of data are required. Furthermore, it allows for working with multi-media data from which results need to be generated and provides features for automated transcriptions.

Synthesis of Results

In order to synthesise the results from the data extraction and to explore relationships between papers and concepts, Petticrew and Roberts (2006) developed a comprehensive list tools and techniques for exploring these relationships, which is provided in Table 6. It is adapted to identify the research publication and their respective relationships.

Table 6

Tools and Techniques for Exploring Relationships

Tools and Techniques	Research Publication
Graphs, frequency distributions, funnel plots, forest plots and L'Abbe plots	
Moderator variables and sub-group analyses	
Idea Webbing and conceptual mapping	
Translation: reciprocal and refutational	
Qualitative case descriptions	
Investigator / methodological triangulation	
Conceptual triangulation	

Note. Tools and techniques for exploring relationships adapted from Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Blackwell Pub.

Collaborative Process

Collaborative processes are concerned with tasks that require interaction of and between multiple researchers, such as the division of tasks and communications. Several approaches have been identified in the literature: Peer review approach (Tsafnat et al., 2014), model driven approach (Barn et al., 2014), visualising relationships (Dobrkovic et al., 2018; Felizardo et al., 2011), symbiotic approaches (Dobrkovic et al., 2018; Goldfarb-Tarrant et al., 2020), crowdsourcing approaches (Mortensen et al., 2017; Weiss, 2016) and many tools incorporate collaborative features such as Covidence, Elicit, Notion and any resource manager with an include and exclude button that can be used by more than one reviewer. Rayyan.ai offers the opportunity to make collaborative systematic reviews, provides training,

priority support and accelerate a review (https://rayyan.ai, retrieved 04.02.2022)https://www.rayyan.ai/sign-up

Peer Review of Proposed Review Protocol

Tsafnat et al. (2014) discuss the peer review approach of a proposed review protocol that can be carried out in order to assess if the SLR process plan is acceptable and adheres to quality standards. Furthermore, the authors describe the creation of a protocol standard which can formally be assessed for consistency, unbiasedness and appropriateness for the research question and as a way of checking the completeness of the inclusion and exclusion criteria.

Collaboration Tools

Barn et al. (2014) used a model driven approach for the development of an open source, web-based, multi-user tool that supports the SLR process for a range of research areas to support collaborative SLRs. The tool supports steps in the SLR process by providing a digital solution for the organisation of and interaction with research articles and allows for collaboration of researchers by having shared access to a project.

The collaborative literature search system (CLSS) has been developed as a solution to create a symbiotic partnership between human researchers and computer agents (Dobrkovic et al., 2018). This system utilises the calculational and computational power of computers to visualise the relationships between scientific research articles in order for researchers to perform qualitative analysis and creative reasoning. Dobrkovic et al. (2018) use seed-based search is used to find related papers by providing a directed graph with metadata of the relationships between research papers.

Crowdsourcing

Crowdsourcing is the distribution of tasks to untrained workers via the Internet. Mortensen et al. (2017) recruited a crowd from Amazon Mechanical Turk which were asked to answer 4 to 5 questions per citation, which were equivalent to the eligibility criteria of the SLR. The answers were aggregated using different algorithms and the outcomes were compared. A reduction in costs for the review of 88% could be achieved, due to the fact that human research experts time costs more than the time of the crowd (Mortensen et al., 2017).

Weiss (2016) describes the concept of crowdsourcing and how it can be leveraged for systematic literature reviews in a newly developed field. The author describes how a group of people can solve complex tasks. In the paper crowdsourcing is viewed as a type of co-creation in which members of the crowd are selected “based on participation criteria, such as,

for their expertise or collaboration history" (Weiss, 2016, p.9). The experts bring insights into which articles can be relevant and consensus is being created through commenting, voting and tagging.

Goldfarb-Tarrant et al. (2020) instructed experts to annotate a given dataset, which was consequently used to train the machine learning algorithm to improve the search process.

Crowdfunding

Crowdfunding for scientific research can be set up to gain exposure, create new collaborations and gain more support for scientists' research. This can be done by providing information about a project and the state of the research in multi-media format.

Software Solutions

A tool called Notion can be used for literature management. The tool includes a glossary to sort papers by topic. Other available features are tables, lists and board view. Many other tools also incorporate collaborative features such as Covidence or Elicit and any resource manager with an include and exclude button that can be used by more than one reviewer.

Channels for Sharing of Publications

Feedback can be gained by sharing a preview with Peers and through the Peer-Review process. Channel for distribution of publication in a field are journals, open-databases, internal databases of research organisations, universities and research associations. Furthermore, links to research publications can be shared via social media, like Twitter, newsletters or blogs.

Current vs. Ideal State – a Comparison

In order to identify areas for improvements in SLRs, an analysis of the current state has been conducted, resulting in numerous approaches that can be utilised for support and aid in the review process.

Current State – Recapitulation of the Literature Search Section

Most researchers do not use any tools to conduct their SLRs (van Altena et al., 2018) except for search and bibliographies. However, a rigid process is most commonly used (Petticrew and Roberts, 2006; Kitchenham and Charters, 2007; Webster and Watson, 2002). Most work is being done manually, which is quite often inefficient, repetitive and prone to errors (Webster and Watson, 2002). Several possible and promising approaches have been

discussed in the literature search section. The following list describes the approaches by using the six dimensions to describe the state-of-the-art practices in SLR.

Search Process

The search process is supported by many tools and approaches. Most commonly used and well-known are the standard search engines of journals or https://scholar.google.com. Many approaches to improve the search using these databases have been proposed, such as the expansion of search terms (Jia & Liu, 2018), using full-text searches (Blanc et al., 2015; Nováček et al., 2009), semantic approaches (Atanassova & Bertin, 2014; Jia & Liu, 2018) the use of indexing mechanisms and annotations (Koukal et al., 2014; Goldfarb-Tarrant et al., 2020; Cohen et al., 2021), bibliometrics (Pulsisri & Vatananan-Thesenvitz, 2018) Checklists and design principles (Bin Ali & Usman, 2018; vom Brocke, 2015; Sturm & Sunyaev, 2019; Schoormann et al., 2018), graphical representations (Watson & Webster, 2020; connectedpapers.com), meta-requirements and meta-search engine (Sturm et al., 2015; Sturm & Sunyaev, 2017), machine learning solutions to support the search process (Marshall & Wallace, 2019; Ros et al., 2017) and other software solutions like elicit.org.

Screening Process

The screening process usually requires the researcher to scan the titles and abstracts to make the inclusion or exclusion decision. However, this process can be supported by machine learning approach for automating text classification (Bekhuis & Demner-Fushman, 2010; Jaspers et al., 2018), automated text annotation approaches (Desclés, 2006), Text mining (Feng et al., 2017; Rathbone et al., 2015), crowdsourcing for citation or abstract screening (Mortensen et al., 2017; Lee et al., 2017; Nama et al., 2019), provision of an overview of the body of knowledge (Pulsisri & Vatananan-Thesenvitz, 2018) and the use of summaries of research articles (Uban and Caragea, 2021; Ibrahim Altmami and El Bachir Menai, 2020; Erera et al., 2019).

Selection Process

The selection process is based on the outcomes of the screening phase. It is the step in which the decision for the inclusion of research publications is made, which in turn are going to be used in the full-text analysis in the data extraction process. The first step is the use of a protocol for article selection (Ferreira et al., 2021) or guidelines (Nama et al., 2019) to add rigor the process. The second step is the evaluation of the quality of review papers using for example benchmarks and quality assurance criteria (Lange & Di lorio, 2014; Leidner, 2018).

Other helpful approaches are the use of concept maps (Felizardo et al., 2017), a node-link graph (Chou & Yang, 2011) or a linked data approach (Tomassetti et al., 2020). Utilising the bibliography for the citation network approach and the theory-driven approach ADIT (Larsen et al., 2019). Furthermore, visual text mining techniques can be used to identify the relevancy of research articles (Felizardo et al., 2012). Moreover, crowdsourcing is a promising approach to article selection, by utilising non-experts in a field to conduct parts of the review process that is very time-consuming (Nama et al., 2019; Lee et al., 2017; Weiss, 2016).

Data Extraction Process

After the selection phase, the full text of all papers that are included are grouped, analysed and relevant information is being collected in order to build the theoretical framework and the literature search section. Consequently, the data are synthesised to answer the research question. Several approaches can aid the process, such as document summarisation (Ibrahim Altmami & El Bachir Menai, 2020; Uban & Caragea, 2021), document classification using machine learning and deep learning methods (Hamad & Saiim, 2014; Rodrigues Alves et al., 2018) or discourse and semantic annotation techniques (Liakata et al., 2012; Bertin, 2014). Furthermore, knowledge graphs can depict relationships between papers (Agrawal, 2021) and other visualisations through contextual and cognitive analysis (Bertin & Atanassova, 2012). In order to save and share annotations, classifications, notes and decisions made during the process, several tools like NVivo or Covidence can be used. Nevertheless, there also exist guidelines for organising and preparing papers for analysis and frameworks (Bandara et al., 2015; Jonnalagadda et al., 2015).

Collaborative Process

Conducting a SLR is usually not a task that one individual researcher can do alone. At least one other researcher is necessary to get an inter-rater reliability score for the inclusion and exclusion of research papers. Most SLR are conducted in a team, for which it is necessary to divide tasks and to share the work that has already been completed, hence to keep every researcher working on the problem in the loop. Many tools exist that can support the researchers in their collaborative actions. These are among many others, Covidence, Elicit, Notion and Rayyan.ai. Barn et al. (2014) developed an open-source collaborative software solution using a model driven approach and Dobrkovic et al. (2018) collaborative literature search system. Furthermore, symbiotic processes between tools, algorithms, experts and researchers can be used to overcome resource scarcity which are described by Goldfarb-Tarrant et al. (2020). Furthermore, crowdsourcing approaches have been utilised by

Mortensen et al. (2017) and Weiss (2016) to outsource some of the tasks of researchers, with astonishing results. Moreover, every paper has to go through the scrutiny of peer review. Therefore, Tsafnat et al. (2014) propose to have peers check and provide feedback on the research paper before publishing it.

Channels for Distribution

Journals are the main research outlets to collect, present and organise research, who mostly have their own, sometimes contrary incentives to open research. Research and profit motives often do not align well.

Desired State

It is interesting to compare the current state to the ideal state, which necessitates to specify the ideal approach to SLRs. The desired situation is that researchers are able to define their research questions, plug them into a software tool and consequently receive a comprehensive and complete list of relevant literature with the visualisation of the relationships between the papers and concepts in that topic, in order for the researchers to be able to immediately start working on answering their research question. Therefore, the aim is to know everything about a certain topic, about methods, about guidelines through a fully automated process which is adapted to the researchers needs.

Discrepancy Between Current and Desired Situation

Every step in the process of SLRs can be subject to automation. However, most approaches try to automate single tasks in conducting the review and do not provide sole solutions to get to the desired state. In this section, the developed dimensions are going to be elaborated on regarding the value of the identified approaches towards to the subject of automation and how promising they are to lead to the desired state.

Procedural Aspects

The creation of the research question is a manual task carried out by the researcher, however tools such as elicit exist that can aid in the development of an appropriate question and can suggest alternative research questions. Furthermore, extensive work has been produced regarding general frameworks for conducting SLR. These works identified that processes, guidelines, protocols and checklists can be subjects for automation. Therefore, a specification manual that is interactable for machines and researchers is the first step towards the desired state.

Search, Screening and Selection Process

All-inclusive databases are necessary to be able to find research articles, which leads to using one search bar instead of many. Semantic approaches, tagging and annotation of research articles in combination with machine learning approaches to create more meta-data of research articles are promising and could lead to the desired state if enough information can be generated and processed according to the aims of the research. In the same vein, full-text search would provide more information that could be used for the improvements in the search process. A combination of these approaches could generate the desired search process and could also automate the screening and selection process, by indicating the relevance of the vast amount of information that are and will be available.

Data Extraction Process

Document summarisation and classification can reduce the time spent trying to identify relevant information from deemed relevant research articles to answer the research question. Moreover, graphical representation of information and relationships between papers can aid in the data extraction process by focussing the researcher's attention to the important information in a research article. Ultimately, natural language models can support this process. However, currently these approaches lack in their accuracy of output. Therefore, the classification of information or more specific of insights can provide researchers with direction, input and knowledge that can be used for example for visualisation purposes. Furthermore, research can be able to interact with the information to judge their relevance or correctness. Ultimately, the trained model could extract the relevant information and write paragraphs of research articles, based on the interaction between information and researchers.

Collaborative Process

Collaboration tools enable researchers to conduct their reviews in teams and to split the tasks at hand. The main issue regarding the desired state is the lack of sharing of the information generated by researchers during the process, which could be used to support other researchers in their manual task, but could also be used to train machine learning models to improve the search, screening and selection processes. Therefore, the main issues are splintered research activities carried out in silos. Moreover, the lack of interactivity with research publications constrains researchers to get a lot of information in an accessible and reliable manner. Moreover, crowdsourcing is an approach that helps to divide up tasks at hand and can be a source for the aforementioned training of the model. Appropriate incentive structures are necessary in order for crowdsourcing to be effective.

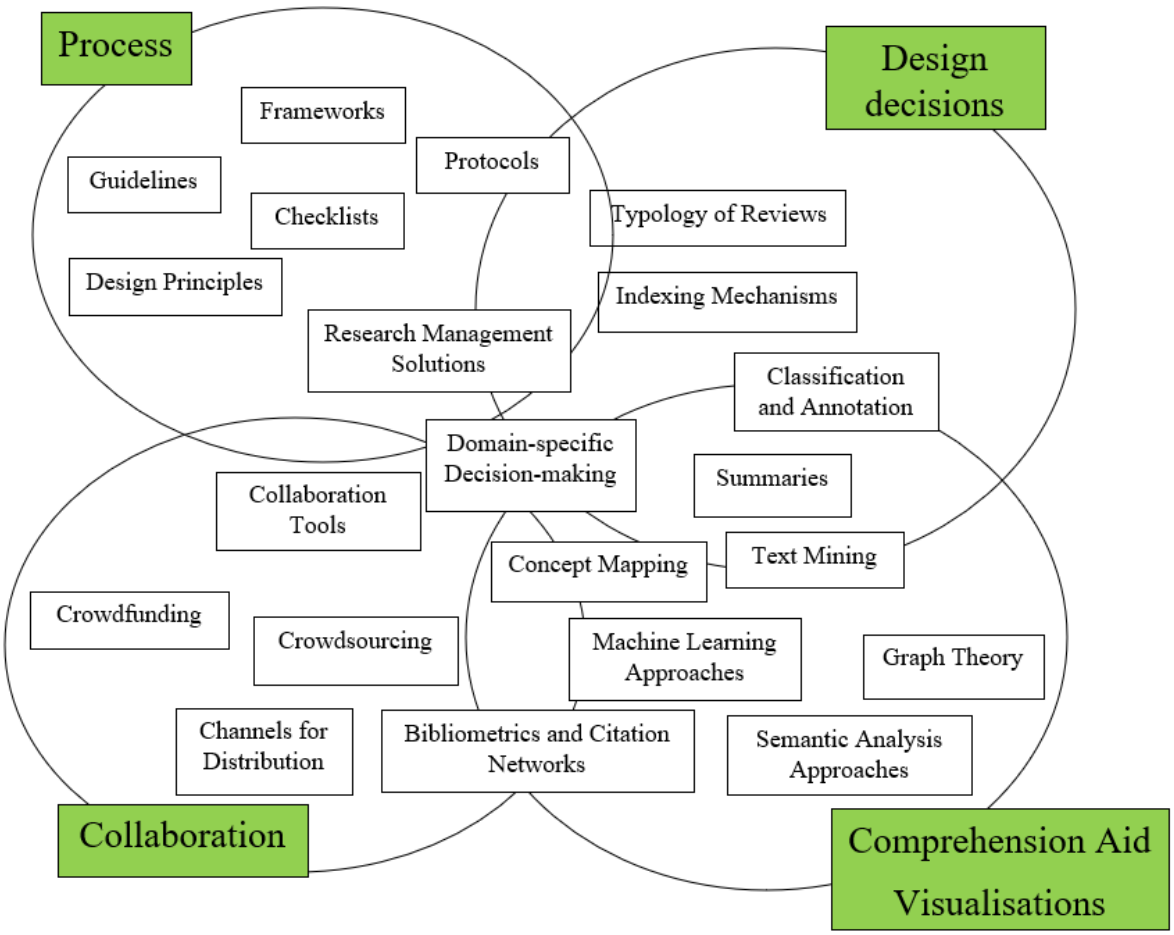
Summary

Overall, the closest to an end-to-end pipeline for the search and selection process has been developed by Goldfarb-Tarrant et al. (2020). The authors use crowdsourcing of experts to annotate research articles, which are used in a machine learning algorithm that sorts the search of research articles in a research field of interest in order for researchers to immediately start with the data extraction process.

Figure 15 illustrates the concepts that have been identified in the literature and provides an overview of possible solutions to tackle issues that arise in the process of conducting systematic literature reviews. The concepts are grouped by similarity and are four general categories have been identified, concerning the process optimisation, design decisions and interactions with research publications, collaboration aspects and understanding, visualisations and data extraction.

Figure 15

Concept Map of Approaches used in the SLR Process Identified in the Literature



Note. The diagram illustrates a concept map of approaches that have been identified in the literature and are grouped by similarity.

Ultimately, a specification manual for the process of SLR that provides the basis for the model, its database and visualisations that is interactable with by researchers and is trained by its interaction with the realm of science is the aim to get to provide the scientific community and beyond with the benefits of up-to-date systematic literature reviews.

Table 7 provides a broad overview of the literature search section. Therefore, it summarises the main findings and links the approaches to the steps in the process and their respective dimensions. Furthermore, it outlines the advantages and disadvantages of the concepts and tools that can be utilised for the automation of SLR.

Table 7

An Overview of Automation Methods Described in the Thesis with their Respective Advantages and Disadvantages

Step of SLR Process	Dimension	Automation method		Advantages		Disadvantages	
		Concepts	Tools	Concepts	Tools	Concepts	Tools
Formulating the Research Question	Procedural Aspects, Collaboration Aspects	Guidelines Checklists, Design Principles, PICO statement (especially for clinical questions)	Elicit	Structured approaches that are defendable, transparent and repeatable	Keywords can be used to provide research questions. Alternative research questions can be generated that are adaptable in their scope.	Takes more time than discussing the RQ with experienced colleagues	Input is needed, hence not fully automated
Design / Plan Review	Procedural Aspects, Collaboration Aspects	Guidelines, Frameworks, Systematic Guides, Protocols, Design Principles, Checklists, Research Management Solutions, PRISMA statement,	Covidence, Elicit, ASReviewer, Rayyan.ai, Buhos, Mendeley, Zotero, Obsidian, Calibre-ebook, EPPI Reviewer, Sesra	Provide structure for the rigorous process, Support decision-making process, Provides replicability, Adaptable to domain specific issues, Less manual errors,	Provision of structure, Provision of features, Provision of task management options, Provision of collaboration opportunities, Saves a lot of time	A lot of preparation and comprehension is necessary to advance on the project	Learning to use tools takes time, Tool uptake is quite slow, 'Valley of Tears'

Conduct the Search	Search Process, Collaboration Aspects	Text similarity, semantic facets, Full-text search, Indexing, Design principles, Machine learning, Text annotation, Text mining, Databases, Encoding core knowledge, Bibliometrics, citation networks, Linked data, Summaries	Elicit, connectedpapers, Research outlets, Literature databases	Reduction of manual errors, Less time and resources necessary, More fine-grained search	Less time and resources necessary, More fine-grained search, more relevant articles, holistic search		Biases in the underlying model of search systems could influence the search
Screen research publications	Screening Process, Collaboration Aspects	Research Management Solutions, Text-classification, Text annotation, Text mining, Benchmarks, Crowdsourcing, Summaries	Covidence, ASReviewer, Elicit	Reduction in resource usage, Outsourcing of specific tasks, More information for decision-making for inclusion and exclusion	Provision of structure and data management, Provision of collaborative features	More information to review for decision-making	Learning to use tools takes time

Appraise the quality and content of research articles	Selection Process, Collaboration Aspects	Visual text mining, concept maps, Protocols, Guidelines, Node-link Graphs, Citation network, Implicit theory, Crowdsourcing, Machine learning, Linked data, Quality evaluation criteria	Transparency, Replicability, Structure, Provision of an overview of the state and further information of the screened research publications	Quality evaluation criteria have to be valid, the optimal level of specificity and optimally peer reviewed

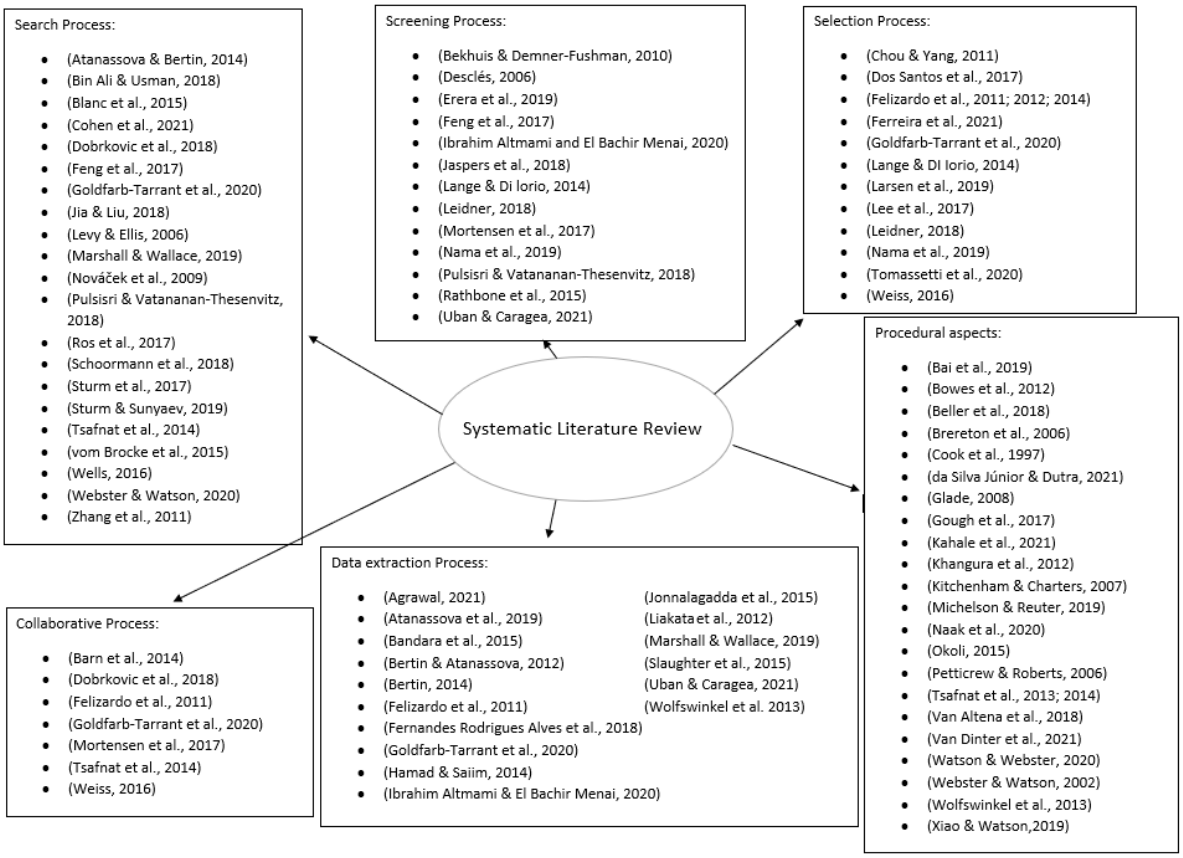
Extract relevant information	Data Extraction Process, Collaboration Aspects	Knowledge graphs, document summarisation, Visualisations, Cognitive Analysis, Deep learning, Classification of full text, Semantic annotation, scientific discourse annotation, Guidelines, Frameworks, Grounded theory, Machine learning classifiers, identification of research objects, updating of reviews	Elicit, NVivo, Notion	Reduction of resources and time, Provision of structure and the ability to save progress, Collaboration opportunities, Sharing of information, More information for improved navigation through research publications, Visualisations to make information more accessible	Structures information, Provision of relevant suggestions, Elicit is able to provide suggested extraction of data	More information to go through, Potential cognitive overload,	An integrated system with all features is not yet available, hence the use of multiple tools is necessary
Reporting on the outcomes of the Literature search	Procedural Aspects, Collaborative Aspects	Guidelines, Frameworks, Checklists, Peer review, Formatting specifications, Domain-specific subtleties, Channels for sharing	MS Word, LaTeX, Twitter, Journals	Highly structured reporting style, Adaptable to the context, Sharing of publications to wider audiences	Word processing tools need to be learned	Reduction in flexibility	Limitations in available features, Training necessary in order to use tools

Note. The table provides an overview of the methods and tools used in each step of the SLR process and their respective dimensions, advantages and disadvantages.

An overview of the research publications included in the review which are ordered by the dimensions they are trying to improve upon is provided in Figure 16. This figure can be used by other scholars to find relevant articles if they are searching for publications in one or more of the dimensions.

Figure 16

Overview of Research Publications Included in the Review and Sorted by their dimension



Note. Overview of all research publications that have been included in the literature review which are sorted by their respective dimension. Some publications appear in more than one dimension.

Concluding Remarks and Discussion

In this study, the scientific literature about systematic literature reviews of the past twenty years has been searched to identify approaches that are directed towards automating steps of systematic literature review processes. 90 publications that capture the strategies and methods for automation have been identified and have been addressed in this thesis. The

choice to adapt the systematic literature review process has been based on the feasibility and the availability of tools that cluster similar publications together, such as connectedpapers.com and elicit. The study has led to an overview of promising approaches for the automation of steps of SLRs and conducting high quality reviews. It discusses the issue of how to approach SLRs in current times and how it could be carried out by aggregating potential combinations of identified approaches by answering the research question:

What methods, approaches and tools can be used to inform the design of improving the effectiveness and efficiency of the search, selection and data extraction process in systematic literature reviews?

Several methods and approaches that improve the processes of SLR have been identified in the six dimensions. Each of the identified approaches leads to improvements in their respective area of focus. The approaches are either of structural or technological nature, hence they bring rigor in the process or provide assistance with technological applications. Each of the identified tools supports some of the manual tasks carried out by researchers, provides assistance in for example managing research publications or by ranking search results to find relevant publications in a shorter period of time. The most promising approach that focusses on the search, screening and selection process utilises expert annotation of publications in a research field, which are used in combination with title and abstract analysis to be used as inputs for a machine learning algorithm. This process provides a reduction in time spend on finding relevant papers for a research question in that area.

The question mentioned in the abstract of "how would the research community and beyond benefit, if the power of SLR could be harnessed at low cost and with high accessibility?" can be answered by utilising the literature search section. It outlines that tool use makes SLR first and foremost much more accessible and appealing, which could lead to an increase in the number of reviews being produced. Furthermore, automation techniques and transparency of processes highlight the importance of a rigorous process in research. Moreover, updatable reviews, also called living reviews, can address the issue of out-of-date publications.

Following the recapitulation of the literature search results and the reconsideration of the desired situation in the light of the results, the concluding section describes the discrepancies between the current and the desired situation, which is the basis for the

implications for further research that could be taken in order to close the gap between the current and the desired situation, hence leading to the automation of SLR.

Ultimately, applying the structure and reporting of SLR improves the repeatability and reliability of research, especially if the steps of the process is reported on, which in turn can then be reviewed, audited and verified and hence, can be recreated. Furthermore, visualisations of topics, concepts, theories, data, methods and their respective relationships and properties can be immensely useful in the decision process of research areas and directions. Moreover, it is also necessary to describe the context and to take contextual variables into account.

Improvements in the Steps of the SLR Process

Automation aspects can support each step in the process of SLR. For each step many approaches, ranging from fully manual to fully automated have been elicited from the literature. The step in which currently has the least automation techniques is the data-extraction process. However, this step could be automated by using natural language models like GPT-3 or in the future GPT-4 to identify relevant sentences from full-text research articles, which are aggregated and can be used by researchers to write their report.

Implications for Practice

This thesis outlines possible solutions to problems that researcher who are conducting a SLR can use to aid in their workflow. Tools, frameworks, guidelines, checklists, and other approaches in the SLR process are explicated and linked to their respective origin. Therefore, researchers are able to get more information about approaches and tools they are interested in. Ultimately, the use of tools and approaches can lead to the reduction of the necessary effort for decision-making in the process and in administrative tasks, which mainly leaves the creative aspects, e.g., the creation of connection and the reporting part.

Limitations

The first limitation is that the literature on which the thesis is based on, might not be as extensive as a complete SLR, due to the fact that in the search process literature has been identified based on a corpus of 19 papers from the pilot study using the aid of tools, such as connectedpapers.com, elicit and web search engines and of search methods, such as forward and backward snowballing. Therefore, some promising approaches might have been missed. During the writing of the report no tools have been used except for the creation of the bibliography. Furthermore, the search process has been supported by tool use. The reason for

combining several tools was that features of tools are very splintered and not bundles in one solution. Another limitation was that only one researcher conducted the main parts of the process, except from one other reviewer who applied the inclusion and exclusion criteria for the inter-rater reliability. Moreover, feedback and discussion have been held with the supervising professor.

Implications for Further Research

This section provides an outlook into what could be accomplished to close the gap between the current and the desired situation.

In the future it would be great to first of all, have the possibilities to interact with research publications, which is not limited to hyperlinks, digital object identifiers and links to profiles of researchers which can be found in publications. It is possible to carry out annotations, meta-data and notes and using research management systems to enrich text. However, these are mostly done in the solitude of research projects. Therefore, sharing, hence increasing the availability of these information is key to the future of SLRs. Moreover, the literature search process needs to be transparent in order to be replicated, hence the search process needs to be tracked automatically and presented in a search protocol which is visible for other researchers.

Second, the cost of high-quality research is very high. If the time being spend on design, data collection, exchange of information with colleagues and synthesis and less on searching for relevant literature would close the gap. Furthermore, SLR could be updated automatically or at least the researcher could be provided with notifications on new publications in their field of study to be included in the update and thus providing the stage for “living reviews”.

Third, the use of natural language processing technologies like the model GPT-3 developed by OpenAI provide the potential to create summaries of scientific articles, to conduct text mining and to aid in the writing process of scientific articles. Therefore, there is need to test the reliability of such summaries by compare them to expert summaries.

Fourth, the role of highlighting of scientific texts needs to be addressed. Each individual with differing degrees of expertise and skills in a specific subject area focusses on different parts of research articles, thus deems different information as relevant. If these highlights and the characteristics of their creator could be aggregated to create a topology of the information in a research article, which can aid in the search process, especially in the

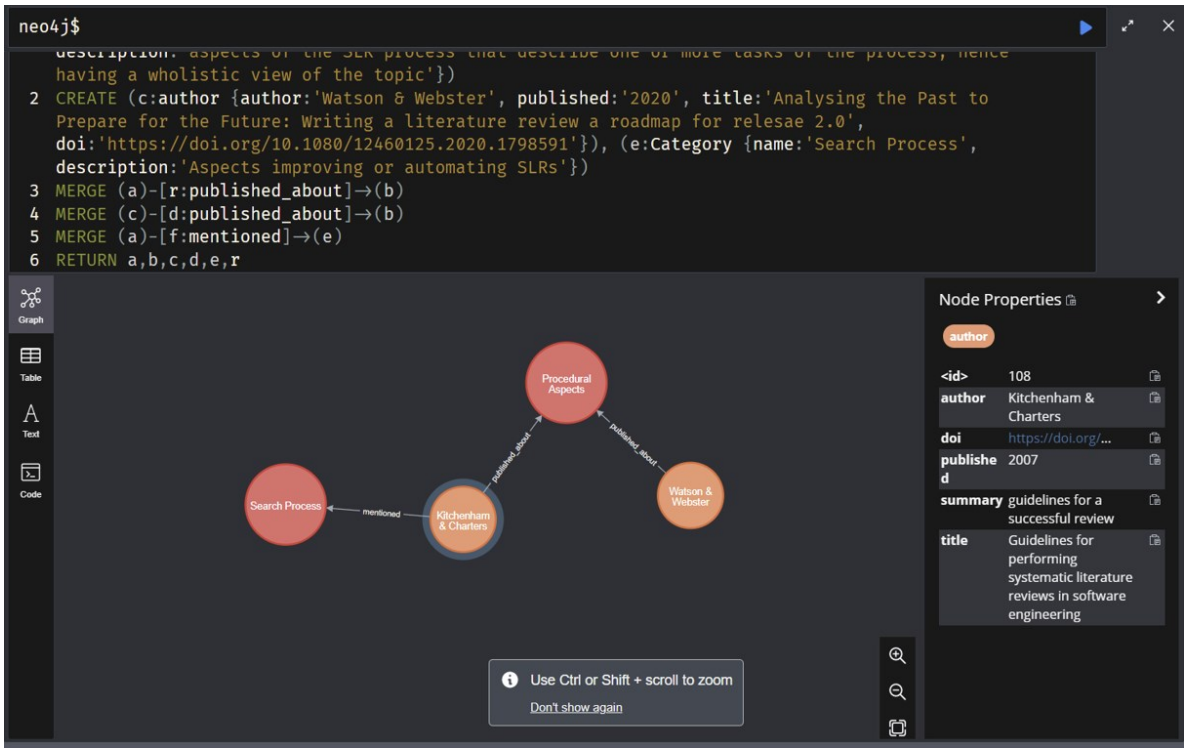
intra-publication search. Furthermore, these data could be utilised to deliver tailored content based on preferences and needs of readers and systematic literature reviewers. It could also enrich the search process by providing more relevant insights into the full-text analysis and be a basis for automatic data extraction.

Fifth, the chances that crowdsourcing can provide has not been used to its full potential and is a key element in the creation of automation of SLR, by carrying out underlying tasks and by adding collaborative decision-making regarding validity and sense-making of information added to a system that provides the necessary features for high quality research. Lastly, the reduction of the necessary effort in administrative tasks and decision-making reduces the strain that SLR have for researchers and would make it more likely for researchers to engage in these activities.

Sixth, the visualisation of the results from the literature search process in an interactive format is possible using a node-link graph. These graphs can show nodes, edges and properties. The software to create these kinds of graphs is called Neo4j, which is able to represent multiple edges compared to other graph software. Each node has adaptable properties which are based on intrinsic information of that node, which can be provided in its the creation and also added later on. Links between the nodes represent their relationship with each other. The underlying data-structure are rows and columns. Figure 17 provides an exemplary representation of the suggestion.

Figure 17

Outline of a Form of Representation of the Information of the review using a Graph



Note. The figure shows an exemplary outline of a node link Graph created with the Neo4j application outlining the potential form of representation of a SLR.

Ultimately, it is necessary to conduct empirical assessment of the use of tools and approaches in order to outline the benefits they can have for researchers and in turn to provide high quality resources and training on how to conduct SLR.

References

Agrawal, K., Mittal, A., & Pudi, V. (2019). Scalable, semi-supervised extraction of structured information from scientific literature. *Proceedings of the Workshop on Extracting Structured Knowledge from Scientific Publications*, 11–20. https://doi.org/10.18653/v1/W19-2602

Ali, N. B., & Usman, M. (2018). Reliability of search in systematic reviews: Towards a quality assessment framework for the automated-search strategy. *Information and Software Technology*, *99*, 133–147. https://doi.org/10.1016/j.infsof.2018.02.002

Allen, I. E. (1999). Estimating time to conduct a meta-analysis from number of citations retrieved. *JAMA: The Journal of the American Medical Association*, *282*(7), 634–635. https://doi.org/10.1001/jama.282.7.634

Altena, A. J., Spijker, R., & Olabarriaga, S. D. (2019). Usage of automation tools in systematic reviews. *Research Synthesis Methods*, *10*(1), 72–82. https://doi.org/10.1002/jrsm.1335

Atanassova, I., & Bertin, M. (2014). Faceted Semantic Search for Scientific Papers. In *Semantic Publishing Challenge co-located with the 11th European Semantic Web Conference ESWC2014*.

Atanassova, I., Bertin, M., & Mayr, P. (2019). Editorial: Mining scientific papers: nlp-enhanced bibliometrics. *Frontiers in Research Metrics and Analytics*, *4*, 2. https://doi.org/10.3389/frma.2019.00002

Bai, Z., Jain, N., & Kurdyukov, R. (2019). Conducting systematic literature reviews in information systems: An analysis of guidelines. *Issues In Information Systems*. https://doi.org/10.48009/3_iis_2019_83-93

Bandara, W., Furtmueller, E., Gorbacheva, E., Miskon, S., & Beekhuyzen, J. (2015). Achieving rigor in literature reviews: Insights from qualitative data analysis and tool-support. *Communications of the Association for Information Systems*, *37*. https://doi.org/10.17705/1CAIS.03708

Barn, B., Raimondi, F., Athiappan, L., & Clark, T. (2014). Slrtool: A tool to support collaborative systematic literature reviews: *Proceedings of the 16th International Conference on Enterprise Information Systems*, 440–447. https://doi.org/10.5220/0004972204400447

Bekhuis, T., & Demner-Fushman, D. (2010). Towards automating the initial screening phase of a systematic review. *MEDINFO 2010*, 146-150.

Beller, E., Clark, J., Tsafnat, G., Adams, C., Diehl, H., Lund, H., Ouzzani, M., Thayer, K., Thomas, J., Turner, T., Xia, J., Robinson, K., & Glasziou, P. (2018). Making progress with the automation of systematic reviews: Principles of the International Collaboration for the Automation of Systematic Reviews (Icasr). *Systematic Reviews*, *7*(1), 77. https://doi.org/10.1186/s13643-018-0740-7

Bertin, M. (2014). Scientific monitoring by mining scientific papers. *Proceedings of the 6th International Conference on Management of Emergent Digital EcoSystems - MEDES '14*, 196–200. https://doi.org/10.1145/2668260.2668306

Bertin, M., & Atanassova, I. (2012). Semantic enrichment of scientific publications and metadata: Citation analysis through contextual and cognitive analysis. *D-Lib Magazine*, *18*(7/8). https://doi.org/10.1045/july2012-bertin

Best qualitative data analysis software for researchers | nvivo. retrieved 10. February 2022, from https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

Blanc, X., Collet, T.-H., Auer, R., Iriarte, P., Krause, J., Légaré, F., Cornuz, J., & Clair, C. (2015). Retrieval of publications addressing shared decision making: An evaluation of full-text searches on medical journal websites. *JMIR Research Protocols*, *4*(2), e38. https://doi.org/10.2196/resprot.3615

Booth, A. (2016). Searching for qualitative research for inclusion in systematic reviews: A structured methodological review. *Systematic Reviews*, *5*(1), 74. https://doi.org/10.1186/s13643-016-0249-x

Bowes, D., Hall, T., & Beecham, S. (2012). SLuRp: a tool to help large complex systematic literature reviews deliver valid and rigorous results. In *Proceedings of the 2nd international workshop on Evidential assessment of software technologies* (pp. 33-36).

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software*, *80*(4), 571–583. https://doi.org/10.1016/j.jss.2006.07.009

Chou, J.-K., & Yang, C.-K. (2011). Papervis: Literature review made easy. *Computer Graphics Forum*, *30*(3), 721–730. https://doi.org/10.1111/j.1467-8659.2011.01921.x

Cohen, A. M., Schneider, J., Fu, Y., McDonagh, M. S., Das, P., Holt, A. W., & Smalheiser, N. R. (2021). *Fifty ways to tag your pubtypes: Multi-tagger, a set of probabilistic publication type and study design taggers to support biomedical indexing and evidence-based medicine* [Preprint]. Health Informatics. https://doi.org/10.1101/2021.07.13.21260468

Connected Papers | Find and explore academic papers. Retrieved 16.11.2021, from https://www.connectedpapers.com/

Cook, D. J. (1997). Systematic reviews: Synthesis of best evidence for clinical decisions. *Annals of Internal Medicine*, *126*(5), 376. https://doi.org/10.7326/0003-4819-126-5-199703010-00006

Covidence—Better systematic review management. Covidence. Retrieved 12.10.2021, from https://www.covidence.org/

Desclés, J. P. (2006). Contextual Exploration Processing for Discourse and Automatic Annotations of Texts. In *FLAIRS Conference* (Vol. 281, p. 284).

Dobrkovic, A., Döppner, D. A., Iacob, M.-E., & van Hillegersberg, J. (2018). Collaborative literature search system: An intelligence amplification method for systematic literature search. In S. Chatterjee, K. Dutta, & R. P. Sundarraj (Hrsg.), *Designing for a Digital and*

Globalized World (Bd. 10844, S. 169–183). Springer International Publishing. https://doi.org/10.1007/978-3-319-91800-6_12

Elicit.org. retrieved 19.02.2022, from https://elicit.org/search

Erera, S., Shmueli-Scheuer, M., Feigenblat, G., Nakash, O. P., Boni, O., Roitman, H., Cohen, D., Weiner, B., Mass, Y., Rivlin, O., Lev, G., Jerbi, A., Herzig, J., Hou, Y., Jochim, C., Gleize, M., Bonin, F., & Konopnicki, D. (2019). A summarization system for scientific documents. *arXiv:1908.11152 [cs]*. http://arxiv.org/abs/1908.11152

Felizardo, K. R., Salleh, N., Martins, R. M., Mendes, E., MacDonell, S. G., & Maldonado, J. C. (2011). Using visual text mining to support the study selection activity in systematic literature reviews. In *2011 international symposium on empirical software engineering and measurement* (pp. 77-86). IEEE.

Felizardo, K.R., Riaz, M., Sulayman, M., Mendes, E., MacDonell, S.G., & Maldonado, J.C. (2011) Analysing the use of graphs to represent the results of systematic reviews in software engineering, in Proceedings of the 25th Brazilian Symposium on Software Engineering (SBES 2011). São Paulo, Brazil, IEEE Computer Society Press, pp.174-183. https://doi.org/10.1109/SBES.2011.9

Felizardo, K. R., Andery, G. F., Paulovich, F. V., Minghim, R., & Maldonado, J. C. (2012). A visual analysis approach to validate the selection review of primary studies in systematic reviews. *Information and Software Technology*, *54*(10), 1079–1091. https://doi.org/10.1016/j.infsof.2012.04.003

Felizardo, K. R., & Carver, J. C. (2020). Automating systematic literature review. In M. Felderer & G. H. Travassos (Hrsg.), *Contemporary Empirical Methods in Software Engineering* (S. 327–355). Springer International Publishing. https://doi.org/10.1007/978-3-030-32489-6_12

Feng, L., Chiam, Y. K., & Lo, S. K. (2017). Text-mining techniques and tools for systematic literature reviews: A systematic literature review. *2017 24th Asia-Pacific Software Engineering Conference (APSEC)*, 41–50. https://doi.org/10.1109/APSEC.2017.10

Fernandes Rodrigues Alves, M., Vasconcelos Ribeiro Galina, S., & Dobelin, S. (2018). Literature on organizational innovation: Past and future. *Innovation & Management Review*, *15*(1), 2–19. https://doi.org/10.1108/INMR-01-2018-001

Ferreira, G. F., Quiles, M. G., Nazaré, T. S., Rezende, S. O., & Demarzo, M. (2021). Automation of article selection process in systematic reviews through artificial neural network modeling and machine learning: Protocol for an article selection model. *JMIR Research Protocols*, *10*(6), e26448. https://doi.org/10.2196/26448

Fink, A. (2020). *Conducting research literature reviews: From the internet to paper* (Fifth edition). Sage.

Glade, M. J. (2008). Systematic Literature Review Specification Manual. *Nutrition*, *24*(5), 495–496. https://doi.org/10.1016/j.nut.2007.12.018

Goldfarb-Tarrant, S., Robertson, A., Lazic, J., Tsouloufi, T., Donnison, L., & Smyth, K. (2020). Scaling systematic literature reviews with machine learning pipelines. *arXiv:2010.04665 [cs]*. http://arxiv.org/abs/2010.04665

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews* (2nd edition). SAGE.

Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, *11*(2), 181–217. https://doi.org/10.1002/jrsm.1378

Hamad, Z., & Salim, N. (2014). Systematic literature review (SLR) automation: a systematic literature review. *Journal of Theoretical and Applied Information Technology*, *59*(3), 661-672.

He Zhang, Babar, M. A., Xu Bai, Juan Li, & Liguo Huang. (2011). An empirical assessment of a systematic search process for systematic reviews. *15th Annual Conference on Evaluation & Assessment in Software Engineering (EASE 2011)*, 56–65. https://doi.org/10.1049/ic.2011.0007

Ibrahim Altmami, N., & El Bachir Menai, M. (2022). Automatic summarization of scientific articles: A survey. *Journal of King Saud University - Computer and Information Sciences*, *34*(4), 1011–1028. https://doi.org/10.1016/j.jksuci.2020.04.020

Jaspers, S., De Troyer, E., & Aerts, M. (2018). Machine learning techniques for the automation of literature reviews and systematic reviews in EFSA. *EFSA Supporting Publications*, *15*(6). https://doi.org/10.2903/sp.efsa.2018.EN-1427

Jonnalagadda, S. R., Goyal, P., & Huffman, M. D. (2015). Automating data extraction in systematic reviews: A systematic review. *Systematic Reviews*, *4*(1), 78. https://doi.org/10.1186/s13643-015-0066-7

Khangura, S., Konnyu, K., Cushman, R., Grimshaw, J., & Moher, D. (2012). Evidence summaries: The evolution of a rapid review approach. *Systematic Reviews*, *1*(1), 10. https://doi.org/10.1186/2046-4053-1-10

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. *EBSE Technical Report,* EBSE-2007-01.

Koukal, A., Gleue, C., & Breitner, M. (2014). Enhancing literature review methods-towards more efficient literature research with latent semantic indexing.

Koukal, A., Gleue, C., & Breitner, M. (2014). *Proceedings of the european conference on information systems (Ecis) 2014, tel aviv, israel, june 9-11, 2014*. http://aisel.aisnet.org/ecis2014/

Lange, C., & Di Iorio, A. (2014). Semantic publishing challenge—Assessing the quality of scientific output. *arXiv:1408.3863 [cs]*. http://arxiv.org/abs/1408.3863

Larsen, K. R., Hovorka, D. S., University of Sydney, Australia, Dennis, A. R., Indiana University, USA, West, J. D., & University of Washington, USA. (2019). Understanding the elephant: The discourse approach to boundary identification and corpus construction for theory review articles. *Journal of the Association for Information Systems*, 887–927. https://doi.org/10.17705/1jais.00556

Lee, Y. J., Arida, J. A., & Donovan, H. S. (2017). The application of crowdsourcing approaches to cancer research: A systematic review. *Cancer Medicine*, *6*(11), 2595–2605. https://doi.org/10.1002/cam4.1165

Leidner, D. (2018). Review and theory symbiosis: An introspective retrospective. *Journal of the Association for Information Systems*, *19*(06), 552–567. https://doi.org/10.17705/1jais.00501

Liakata, M., Thompson, P., de Waard, A., Nawaz, R., Pander Maat, H., & Ananiadou, S. (2012, July). A three-way perspective on scientific discourse annotation for knowledge extraction. In *Proceedings of the Workshop on Detecting Structure in Scholarly Discourse* (pp. 37-46). The Association for Computational Linguistics (ACL).

Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science*, *9*.

Marshall, I. J., & Wallace, B. C. (2019). Toward systematic review automation: A practical guide to using machine learning tools in research synthesis. *Systematic Reviews*, *8*(1), 163, s13643-019-1074–1079. https://doi.org/10.1186/s13643-019-1074-9

Michelson, M., & Reuter, K. (2019). The significant cost of systematic reviews and meta-analyses: A call for greater involvement of machine learning to assess the promise of clinical trials. *Contemporary Clinical Trials Communications*, *16*, 100443. https://doi.org/10.1016/j.conctc.2019.100443

Molléri, J. S., da Silva, L. E., & Benitti, F. B. V. (2013). Proposal of an Automated Approach to Support the Systematic Review of Literature Process (S). In *SEKE* (pp. 488-493).

Mortensen, M. L., Adam, G. P., Trikalinos, T. A., Kraska, T., & Wallace, B. C. (2017). An exploration of crowdsourcing citation screening for systematic reviews. *Research Synthesis Methods*, *8*(3), 366–386. https://doi.org/10.1002/jrsm.1252

Naak, A., Hage, H., & Aïmeur, E. (2008, July). Papyres: a research paper management system. In *2008 10th IEEE Conference on E-Commerce Technology and the Fifth IEEE Conference on Enterprise Computing, E-Commerce and E-Services* (pp. 201-208). IEEE.

Nama, N., Sampson, M., Barrowman, N., Sandarage, R., Menon, K., Macartney, G., Murto, K., Vaccani, J.-P., Katz, S., Zemek, R., Nasr, A., & McNally, J. D. (2019). Crowdsourcing the citation screening process for systematic reviews: Validation study. *Journal of Medical Internet Research*, *21*(4), e12953. https://doi.org/10.2196/12953

Novacek, V., Groza, T., Handschuh, S., & Decker, S. (2010). Coraal—Dive into publications, bathe in the knowledge. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3199470

O'Connor, A. M., Tsafnat, G., Gilbert, S. B., Thayer, K. A., & Wolfe, M. S. (2018). Moving toward the automation of the systematic review process: A summary of discussions at the second meeting of International Collaboration for Automation of Systematic Reviews (ICASR). *Systematic Reviews, 7*(1), 3. https://doi.org/10.1186/s13643-017-0667-4

Okoli, C. (2012). *A critical realist guide to developing theory with systematic literature reviews* (SSRN Scholarly Paper ID 2115818). Social Science Research Network. https://papers.ssrn.com/abstract=2115818

Okoli, C. (2015). A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems*, *37*. https://hal.archives-ouvertes.fr/hal-01574600

Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1954824

Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, *52*(2), 183–199. https://doi.org/10.1016/j.im.2014.08.008

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Blackwell Pub.

Pulsiri, N., & Vatananan-Thesenvitz, R. (2018). Improving systematic literature review with automation and bibliometrics. *2018 Portland International Conference on Management of Engineering and Technology (PICMET)*, 1–8. https://doi.org/10.23919/PICMET.2018.8481746

Ranard, B., Ha, Y., Meisel, Z., Asch, D., Hill, S., & Becker, L. (o. J.). Crowdsourcing—Harnessing the Masses to Advance Health and Medicine, a Systematic Review. *Journal of Global Health*, *7*(2), 020602. https://doi.org/10.7189/jogh.07.020601

Rathbone, J., Hoffmann, T., & Glasziou, P. (2015). Faster title and abstract screening? Evaluating Abstrackr, a semi-automated online screening program for systematic reviewers. *Systematic Reviews*, *4*(1), 80. https://doi.org/10.1186/s13643-015-0067-6

Ros, R., Bjarnason, E., & Runeson, P. (2017). A machine learning approach for semi-automated search and selection in literature studies. *Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering*, 118–127. https://doi.org/10.1145/3084226.3084243

Rousseau, D. M., Manning, J., & Denyer, D. (2008). Evidence in management and organizational science: Assembling the field's full weight of scientific knowledge through syntheses. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1309606

Santos, V. dos, Souza, É. F. de, Felizardo, K. R., & Vijaykumar, N. L. (2017). Analyzing the use of concept maps in computer science: A systematic mapping study. *Informatics in Education*, *16*(2), 257–288. https://doi.org/10.15388/infedu.2017.13

Schoormann, T., Behrens, D., Fellmann, M., & Knackstedt, R. (2018). Design principles for supporting rigorous search strategies in literature reviews. *2018 IEEE 20th Conference on Business Informatics (CBI)*, 99–108. https://doi.org/10.1109/CBI.2018.00020

Silva Júnior, E. M. da, & Dutra, M. L. (2021). A roadmap toward the automatic composition of systematic literature reviews. *Iberoamerican Journal of Science Measurement and Communication*, *1*(2), 1–22. https://doi.org/10.47909/ijsmc.52

Slaughter, L., Berntsen, C. F., Brandt, L., & Mavergames, C. (2015). Enabling living systematic reviews and clinical guidelines through semantic technologies. *D-Lib Magazine*, *21*(1/2). https://doi.org/10.1045/january2015-slaughter

Staples, M., & Niazi, M. (2007). Experiences using systematic review guidelines. *Journal of Systems and Software*, *80*(9), 1425-1437.

Sturm, B., Schneider, S., & Sunyaev, A. (2015). Leave No Stone Unturned: Introducing a Revolutionary Meta-search Tool for Rigorous and Efficient Systematic Literature Searches. In *ECIS*.

Sturm, B., & Sunyaev, A. (2017). You Can't Make Bricks Without Straw: Designing Systematic Literature Search Systems. In *ICIS*.

Sturm, B., & Sunyaev, A. (2019). Design principles for systematic search systems: A holistic synthesis of a rigorous multi-cycle design science research journey. *Business & Information Systems Engineering*, *61*(1), 91–111. https://doi.org/10.1007/s12599-018-0569-6

Templier, M., & Paré, G. (2015). A framework for guiding and evaluating literature reviews. *Communications of the Association for Information Systems*, *37*(1). https://doi.org/10.17705/1CAIS.03706

Tomassetti, F., Rizzo, G., Vetro, A., Ardito, L., Torchiano, M., & Morisio, M. (2011). Linked data approach for selection process automation in systematic reviews. *15th Annual Conference on Evaluation & Assessment in Software Engineering (EASE 2011)*, 31–35. https://doi.org/10.1049/ic.2011.0004

Tsafnat, G., Dunn, A., Glasziou, P., & Coiera, E. (2013). The automation of systematic reviews. *BMJ*, *346*(jan10 1), f139–f139. https://doi.org/10.1136/bmj.f139

Tsafnat, G., Glasziou, P., Choong, M. K., Dunn, A., Galgani, F., & Coiera, E. (2014). Systematic review automation technologies. *Systematic Reviews*, *3*(1), 74. https://doi.org/10.1186/2046-4053-3-74

Uban, A. S., & Caragea, C. (2021). Generating summaries for scientific paper review. *arXiv:2109.14059 [cs]*. http://arxiv.org/abs/2109.14059

van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdema, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Hindriks, S., Tummers, L., & Oberski, D. L. (2021). An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence*, *3*(2), 125–133. https://doi.org/10.1038/s42256-020-00287-7

van Dinter, R., Tekinerdogan, B., & Catal, C. (2021). Automation of systematic literature reviews: A systematic literature review. *Information and Software Technology*, *136*, 106589. https://doi.org/10.1016/j.infsof.2021.106589

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in

information systems research. *Communications of the Association for Information Systems*, *37*. https://doi.org/10.17705/1CAIS.03709

Watson, R. T., & Webster, J. (2020). Analysing the past to prepare for the future: Writing a literature review a roadmap for release 2.0. *Journal of Decision Systems*, *29*(3), 129–147. https://doi.org/10.1080/12460125.2020.1798591

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, xiii-xxiii.

Weiss, M. (2016). Crowdsourcing literature reviews in new domains. *Technology Innovation Management Review*, *6*(2), 5–14. https://doi.org/10.22215/timreview/963

Wells, D. (2016). Library discovery systems and their users: A case study from curtin university library. *Australian Academic & Research Libraries*, *47*(2), 92–105. https://doi.org/10.1080/00048623.2016.1187249

Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C. P. M. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, *22*(1), 45–55. https://doi.org/10.1057/ejis.2011.51

Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research*, *39*(1), 93–112. https://doi.org/10.1177/0739456X17723971