USING SUFFIX TREE CLUSTERING METHOD TO SUPPORT THE PLANNING PHASE OF SYSTEMATIC LITERATURE REVIEW

Luyi Feng[1]*, Yin Kia Chiam*[2]*, Erma Rahayu binti Mohd Faizal Abdullah*[3]*, Unaizah Hanum Obaidellah*[4]

[1,2,3,4]Faculty of Computer Science and Information Technology
University of Malaya
50603 Kuala Lumpur
Malaysia

Email: janqualine@163.com[1], yinkia@um.edu.my[2], erma@um.edu.my[3], unaizah@um.edu.my[4]

ABSTRACT

Systematic Literature Review (SLR) is an information-based process throughout which each stage must be carefully and systematically designed in planning phase. Important decisions need to be made about the various choices involved in the SLR planning phase. However, due to the necessarily comprehensive and rigorous nature of SLR, many researchers have difficulties in planning SLRs sufficiently and effectively. In recent decades, the use of text mining (TM) techniques has been introduced to facilitate SLR process, especially at conducting review phase to support the searching and selection of studies. However, so far, TM techniques have not been applied to support any activities during SLR planning phase. In this research, we proposed a method that apply Suffix Tree Clustering (STC) method to support decision-making activities within the SLR planning phase. This method also known as SLR Planning Based on Suffix Tree Clustering (SLRP-STC) method. It aims to help reviewers in three ways: 1) to extract research topics easily; 2) to better interpret extracted information; and 3) to quickly realize and refine a proposed search string that is poorly formulated or inappropriate. A case study was conducted by comparing the proposed method with manual approach. It is observed that the use of SLRP-STC method can improve the planning of SLR.

Keywords: Systematic Literature Review (SLR), planning phase; decision-making, Suffix Tree Clustering (STC), text mining (TM)

1.0 INTRODUCTION

Systematic Literature Review (SLR), or Systematic Review (SR), provides a fair and reliable approach to identify, evaluate, and synthesize all primary studies, in attempt to address specific research questions [1]. SLR differs from ordinary review process in term of formal planning. Staples and Niazi [2] define a well-planned SLR as “independently replicable and of greater scientific value than ordinary literature reviews”. For another, creating a comprehensive review plan is the key to the success of SLR as it drives the entire review methodology. Effective SLR planning takes into consideration all aspects of planning including identification of review need, specification of research questions, development, and validation of the review protocol.

Performing SLR comprises three main phases: 1) planning the review, 2) conducting the review, and 3) reporting the review [5]. Particularly, planning phase is important to ensure the overall quality of SLR process. During the planning phase, reviewers are required to have adequate knowledge of the proposed topic area so they are likely to make a strategic plan for the SLR. For example, reviewers should confirm that there are likely to be sufficient relevant studies to make their SLRs worthwhile. They should also obtain the necessary information relevant to the studies, such as the number of independent studies that make review worthwhile, the research topics, the commonly used terms, and keywords. In this case, effective knowledge acquisition and discovery process can be useful and necessary in supporting informed decisions for SLR planning. The emergence of text mining (TM) techniques offers one possible way to facilitate the SLR process and thus reduce the workload of performing SLRs [3]. The primary goal of TM is to support the data exploratory and data analysis by extracting useful information from unstructured textual documents and presenting the distilled knowledge in a concise form [4].

In practice, many SLRs need to be iterative because review plans keep changing when researchers obtain better idea of the literature, such practice is extremely laborious and inefficient [6]. One way of addressing these difficulties is to have extensive SLR planning activities before the conducting phase. It is believed that adequate planning activities is likely to improve the overall performance of the SLR process [7]. Poor planning and inappropriate review strategy will affect the overall quality and efficiency of SLR process. From another perspective, there is a research gap lies in proposing specific TM techniques to support the SLR planning phase. To the best of our knowledge, the majority studies on supporting SLRs primarily focus on applying TM techniques in SLR conducting phases, especially at the stages of searching and selecting primary studies. None of them have ever employed such techniques in SLR planning phase. Based on our findings from the SLR [15], TM techniques and tools used to support SLR process do not address the challenges faced by researchers during the SLR planning phase. Indexing of documents are important to present the content of the documents [16].

Given the above, the main goal of this work is to propose a new method, SLRP-STC (Systematic Literature Review Planning Based On Suffix Tree Clustering) method to support effective knowledge acquisition and knowledge discovery process involved in SLR planning phase. To evaluate our approach, we have conducted a case study to compare the use of SLRP-STC method with the traditional, manual approach in planning a SLR. Our main contributions are encouraging the adoption of SLR methodology in the software engineering (SE) field by facilitating the SLR planning process. In addition, SLRP-STC contributes to support the data exploratory and data analysis involved during the SLR planning phase by clustering and mining the Web retrieved documents automatically. This can help researchers and students to extract research topics easily, to better interpret extracted information and to quickly realize the proposed search string is poorly formulated or inappropriate, and refine the search strings.

The remainder of this paper is organized as follows: Section 2 presents background and overview of related work. Section 3 explains the SLRP-STC method in details. Section 4 describes the case study used to evaluate our proposed method. Section 5 presents a brief discussion of our findings. Finally, conclusion and future work are presented in Section 6.

2.0 BACKGROUND AND RELATED WORK

This section presents brief background information on SLR planning phase, existing applications of TM techniques to support SLR process and an overview of STC method.

2.1 SLR Planning Phase

With the purpose to appropriately define the entire review strategy, the SLR planning phase is mainly concerned with integrating adequate knowledge of proposed topic area that enables reviewers to make thoughtful decisions involving planning phase. The decision-making activities with the planning phase are illustrated in Fig. 1.

A SLR typically starts with justifying the genuine need of undertaking the review, such as the likely existence of a sufficient number of independent studies that have not been previously organized. The number of studies that makes SLR worthwhile depends on the topic area of interest and should be confirmed by review team. Reviewers should also determine whether SLR studies already exist in the topic of interest. If existing reviews do not cover the proposed topic of interest, or those that do exist are out of date, researchers are encouraged to continue planning the SLR. Otherwise, reviewers should carefully analyze these reviews and decide whether the review results can be directly adopted or need to be updated.

In order to develop review protocol, research questions need to be well defined and agreed before executing SLR process. Specifying appropriate research questions is critical because the questions are used to derive the entire SLR methodology. Reviewers should quickly identify a broad topic area and classify the specific topics studied in the broad topic area. Classification of topics is used to identify boundaries and context for the review. It also helps to interpret the identified studies and to narrow down the review scope.

A protocol specifies the overall SLR approach that will be used to conduct and report the review phase. The major decision-making problem with protocol development is to devise an appropriate search strategy. The difficulty lies in searching for topics that are unlikely to be the main research topic of research papers.

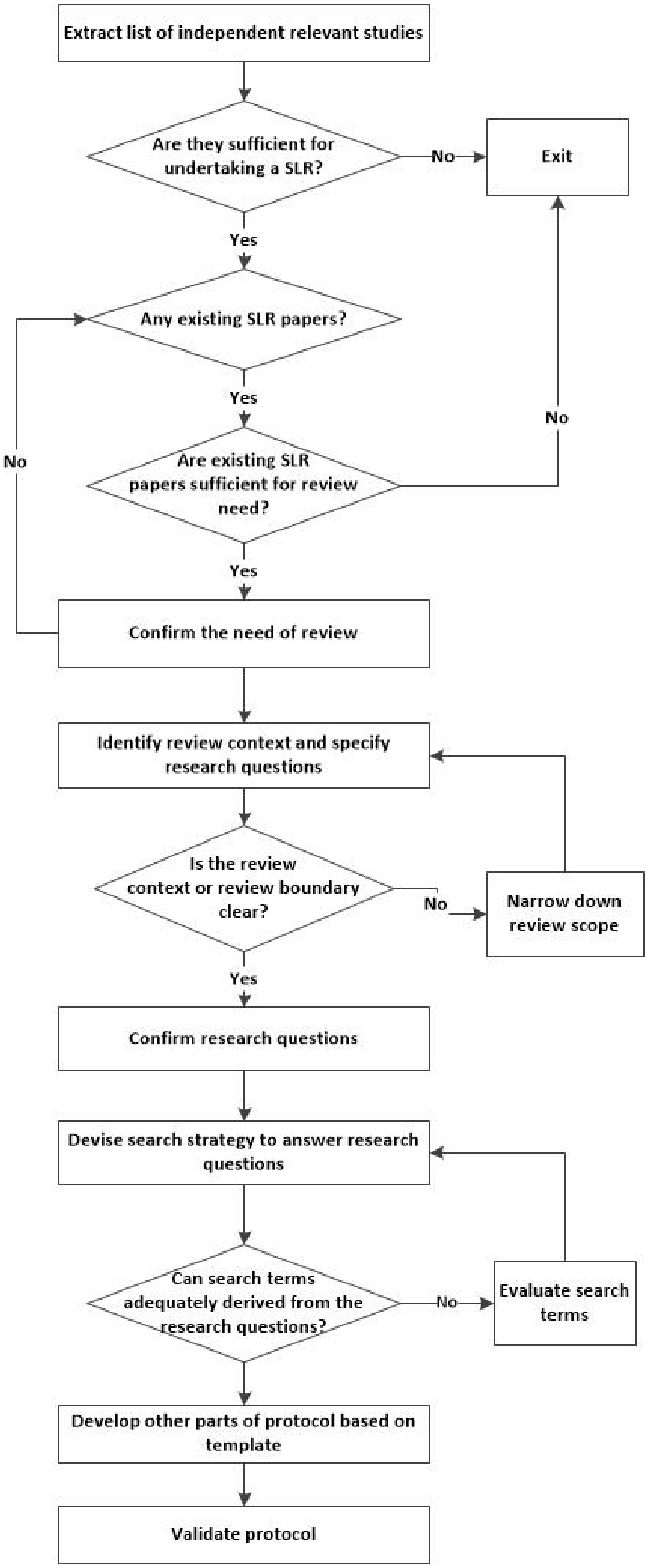

Extract list of independent relevant studies
Are they sufficient for undertaking a SLR?
No
Exit
Yes
Any existing SLR papers?
Yes
No
No
Are existing SLR papers sufficient for review need?
Yes
Confirm the need of review
Identify review context and specify research questions
Is the review context or review boundary clear?
No
Narrow down review scope
Yes
Confirm research questions
Devise search strategy to answer research questions
Can search terms adequately derived from the research questions?
No
Evaluate search terms
Develop other parts of protocol based on template
Validate protocol

Fig. 1: Decision making activities involved in SLR planning phase (adapted from [17])

The reasons account for the problems are (1) the specific research approaches are seldom identified in the title, abstract and keywords of primary studies, and (2) SE researchers are extremely poor at correctly specifying the approaches they used. Thus, reviewers need to perform relatively broad search to identify a known set of papers to assess the search strategy. This prepares the reviewers to manage a large number of candidate primary studies of the papers within the broad topic area. In this case, identification of commonly used terms covered by the studies can be used to devise automated search strings. Due to the decision-making activities in protocol validation highly rely on the expertise and domain knowledge of reviewers, the last stage of the SLR planning phase will not be supported by the proposed method in this research.

2.2 Existing Applications of TM Techniques to Support SLR Process

TM, also referred as intelligent text analytics, text data mining, and knowledge discovery in text, is the process of deriving useful information from text. The process often involves various techniques, including data mining, machine learning, natural language processing (NLP), and knowledge management [8]. TM techniques works by converting unstructured data into structured data. Typical TM applications includes document classification, document clustering, information extraction, information visualization, information retrieval, and document summarization [9]. Various studies [9-14] have discussed applications of TM techniques.

In our previous study [15], four main applications of TM techniques have been identified through a review process: 1) visual text mining (VTM), 2) federated search strategy, 3) automated document/text classification using machine learning (ML) techniques, and 4) document summarization using ML techniques. Table 1 shows a list of eight studies in software engineering domain that have applied TM techniques, especially for information visualization and clustering techniques to support the SLR process. These studies are journal articles, symposium and conference papers, book sections, and workshop papers published in SE discipline. Majority of them focus on building machine learning classifiers for automatic study selection. As shown in Table 1, tools have been developed or used to support their proposed methods.

Table 1: Selected primary studies that using TM techniques to support the SLR process

ID	Authors	Year	Ariticle Type	TM Technique(s) used	Tool(s)
S01	Malheiros et al. [18]	2007	Symposium	VTM	PEx
S02	Felizardo et al. [19]	2007	Conference	VTM	PEx
S03	Fernández-Sáez et al. [20]	2010	Conference	The authors did not specify which TM technique was used to cluster the documents	SLR-Tool*
S04	Tomassetti et al. [21]	2011	Conference	Naïve Bayes Classifier	DBPedia
S05	Ghafari et al. [22]	2012	Journal	Federated Search Model	Federated Search tool*; Nvivo, Zotero
S06	Fabrri et al. [23]	2012	Journal	Vector Processing Model	StArT*
S07	Torres et al. [24]	2013	Workshop	Rule-based and ML algorithm (Nearest Neighbor Heuristic)	A tool based on Weka library*
S08	Felizardo et al. [25]	2014	Conference	VTM, KNN (K-Nearest Neighbor) Edges Connection Technique	ReViS*

*Note: * = Specific SLR Tool developed by the researchers*

2.3 Overview of the STC Method

In this research, we applied the STC Method to support the planning of SLRs. This method allows interactive data exploration of text documents in the Web environment [26]. It can be viewed as a dynamic VTM method that combines document clustering and information retrieval techniques. STC method was first proposed by Gusfield [26]. The author proposed a phrase-based method for automatic clustering search results based on STC method. The STC method is designed to identify phrases that are common in the document set and uses these phrases as the basis for creating clusters [27]. One feature that distinguishes STC from other clustering methods, is it treats documents as a string, rather than bag of words. STC method uses a suffix tree to identify Web documents that share common phrases and automatically classify these documents in descriptive categories. STC method [27] has four logical steps: (1) Retrieve web documents, (2) Clean documents, (3) Construct Suffix Tree, and (4) Merge Clusters. The first two steps are about crawling and parsing document collections. In the third step, a suffix tree from all sentences of documents is built incrementally. In the final step, each base cluster is weighted and merged with high degree of overlap in their document collections.

The differences between STC methods and typical VTM approaches are shown as follows.

- VTM techniques cluster local, off-line documents collection that has standardized format, whereas STC techniques retrieve online documents and cluster Web search results.
- VTM techniques calculate document similarity and use multidimensional projection techniques to place document on a 2D visual map, using hierarchy, model-based clustering algorithms (*viz.*, k-means, kNN). And STC method applies incremental clustering algorithm.

Considering the literature is frequently updated, it is easier to collect, cluster, and analyze documents in dynamic Web environment. In this case, STC method can be useful in supporting SLR planning phase in three ways: (a) to quickly identify research topics; (b) to help them better interpret extracted information; (c) to realize early that queries are poorly formulated (either too general or inappropriate) and hence refine them.

3.0 SLRP-STC METHOD

SLR is primarily an information-based process where each stage must be fairly and systematically designed during planning phase. Due to the comprehensive and rigorous nature of the SLR, exhaustive planning activities are required. In order to guarantee the effectiveness and robustness of SLR process, we propose the SLRP-STC method, to support informed decision-making in the SLR planning phase. The following sub-sections describe the proposed method.

3.1 Overview of SLRP-STC Method

As shown in Fig. 2, SLRP-STC method supports the three main stages of SLR planning: (1) justifying the need for a systematic review or mapping study, (2) specifying research questions, and (3) developing review protocol (i.e. focus on devising automated search strings). This method intends to address the problems faced in the three aforementioned stages by applying the STC method to retrieve and cluster the web documents. STC method consists of four main logical steps: (1) Retrieve web documents, (2) Clean documents, (3) Identify Phrase Clusters, and (4) Merge Clusters. In this work, we added an additional step, (5) Study Visualization to provide the visual representations of the relevant article information extracted and analyzed from the web documents.

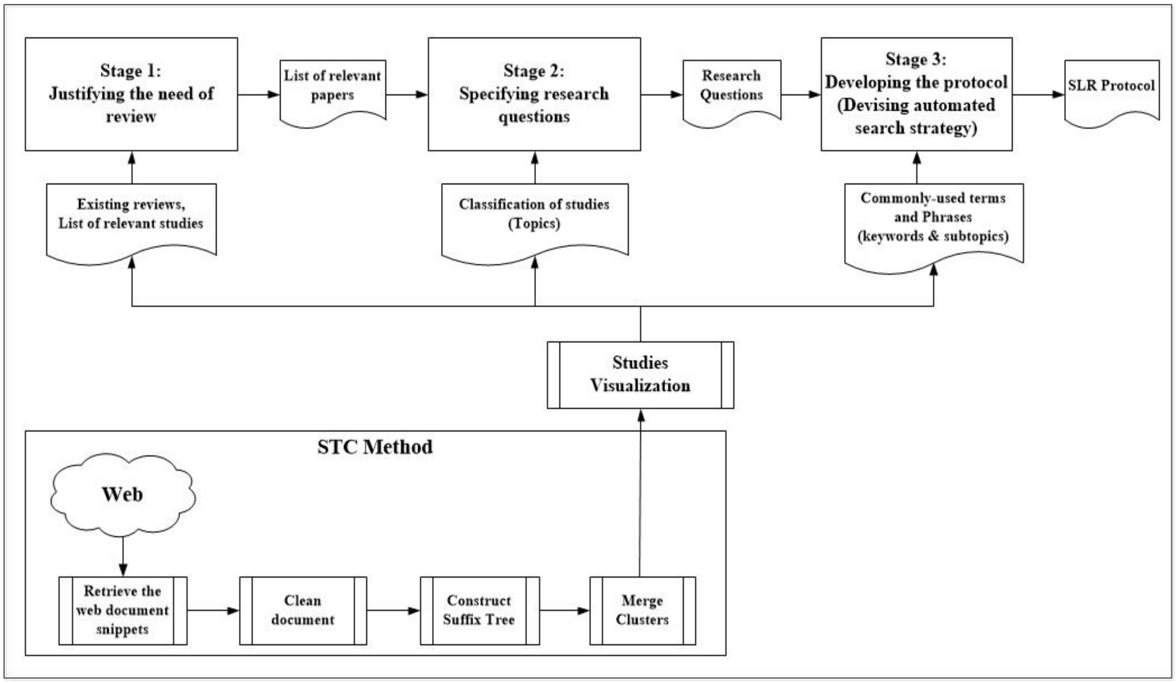

Stage 1:
Justifying the need of review
List of relevant papers
Stage 2:
Specifying research questions
Research Questions
Stage 3:
Developing the protocol (Devising automated search strategy)
SLR Protocol
Existing reviews, List of relevant studies
Classification of studies (Topics)
Commonly-used terms and Phrases (keywords & subtopics)
Studies Visualization
STC Method
Web
Retrieve the web document snippets
Clean document
Construct Suffix Tree
Merge Clusters

Fig. 2: Overview of SLRP-STC Method

3.1.1 Stage 1: Justifying the Need of Review

The decision-making activities with this process are illustrated in Fig. 3. In order to help reviewers quickly justify the need of undertaking SLR, our method is designed to facilitate the knowledge acquisition process for planning phase. A federated search method is offered to help the reviewers to easily extract the necessary knowledge of a proposed topic, where it retrieves the relevant Web documents from different data sources. This method helps the reviewers to quickly obtain the knowledge about existing studies, including a list of relevant studies, and the number of all relevant studies regarding to the proposed topic. Reviewers can also specify the search options to seek for existing reviews within the topic of interest. Next, reviewers briefly construct the most relevant studies to have some background knowledge of the proposed topic area. This enables them to determine whether they should proceed with the review in an objective and fair manner.

Stage 1 (Justifying the need for a systematic review) consists of the following steps:

1. Checking whether any systematic reviews or mapping studies already exist in the topic area.
2. In the case of an outdated systematic review, identifying any limitations with the initial protocol and amend the process. A list of primary studies are identified from the existing reviews. The list will become the basis set of known papers that can be used to define and revise search strings for digital libraries.
3. If there are no previous reviews, identifying a list of primary studies to make sure that there are likely to be sufficient relevant papers to make a systematic review worthwhile. Identify the number of independent studies by conducting a quick informal search using Google Scholar or a digital indexing system to look for relevant studies.
4. Finally, justifying the need for a review based on the information collected in Steps 1, 2 and 3.

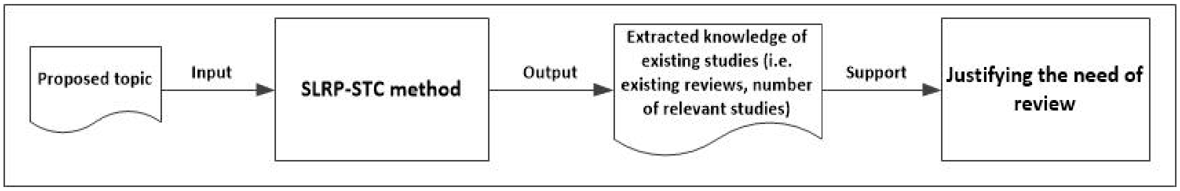

Proposed topic
Input
SLRP-STC method
Output
Extracted knowledge of existing studies (i.e. existing reviews, number of relevant studies)
Support
Justifying the need of review

Fig. 3: Identifying the need of review in Stage 1

3.1.2 Stage 2: Specifying Research Questions

In this stage, the SLRP-STC method aims to support the reviewers to quickly identify the research context and specify the research questions. Fig. 4 shows the decision-making activities involved in Stage 2. For SLRs, research questions need to be well defined and agreed before the review protocol is developed. Specifying appropriate research questions is one of the most critical components of SLR protocol, because the research questions are used to devise other protocol elements, and to derive the entire SLR methodology. Reviewers should quickly identify a broad topic area and classify the specific topics studied in the broad topic area.

Clustering techniques have been used to extract topics from Web documents by grouping similar documents. The classification of topics is useful and necessary in helping reviewers to identify review context and specifying research questions. Moreover, clustering is important in data exploratory and data analysis. In this research, we apply the STC algorithm to cluster retrieved Web documents and to generate classification of topics. In classical algorithms, similarity calculation is usually represented by numerical data, such as distance. The main problem with classical algorithms like HAC, and k-means, is that they are very sensitive to the halting criterion. This sensitivity often causes poor clustering performance in the Web environment where the query results could extremely vary.

Stage 2 (Specifying research questions) comprises the following steps:

1. Identifying a broad topic area and classifying the specific topics studied in the broad topic area.
2. Classifying topics to identify boundaries and context for the review.
3. Interpreting the identified studies and to narrow down the review scope.
4. Defining question elements based on PICO criteria [29]:
   - Population (P): A question which may refer to very specific population groups (e.g. a specific SE role, a category of software engineer, an application area, or an industry group).
   - Intervention (I): The software engineering methodology, tool, technology or procedure that address a specific issue (e.g. technologies to perform specific tasks).
   - Comparison (C): The comparison of software engineering methodology, tool, technology or procedure Outcomes: prediction or estimate accuracy.
   - Outcomes (O): Factors of importance to practitioners (e.g. improved reliability, reduced production costs, and reduced time to market)
5. Formulating research questions for SLR.

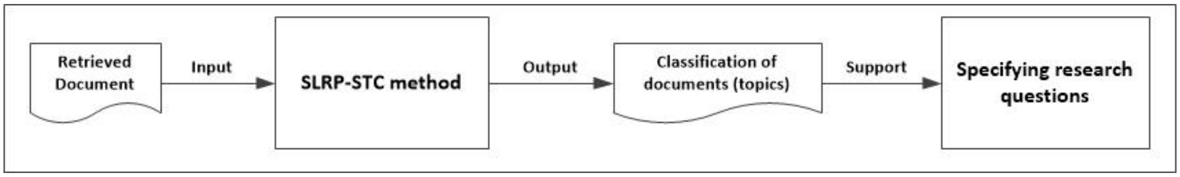

Retrieved Document
Input
SLRP-STC method
Output
Classification of documents (topics)
Support
Specifying research questions

Fig. 4: Specifying research questions in Stage 2

3.1.3 Stage 3: Developing Review Protocol

Devising search strategy is of great importance because the rigor of the search process is one of the main factors that differentiates SLR from other literature review process. However, generating a good automated search strategy is a challenging task. Kitchenham and Brereton [30] suggest reviewers to use different combinations of search terms derived from the research questions. Such practice typically generates thousands, if not millions of search results (citations) that are irrelevant. Additionally, there would be lots of synonyms and abbreviations for specific topic keyword, as not all search terms that make up of the combination are necessary and useful. Using inappropriate search terms would significantly increase the review workload and distract reviewers from the most relevant studies. Fig. 5 shows the decision-making activities involved in Stage 3. The proposed method helps the reviewers to do a relatively broad search from different data sources (based on a federated search method) to identify a known set of papers to assess search strategy. Terms and phrases are identified and ranked according to their significance (weighted score). The core mining operation follows the procedure of STC method (see Section 3.2). Moreover, the identified keywords and subtopics can be used to help reviewers to justify the validity and the accuracy of the proposed automated search strategy.

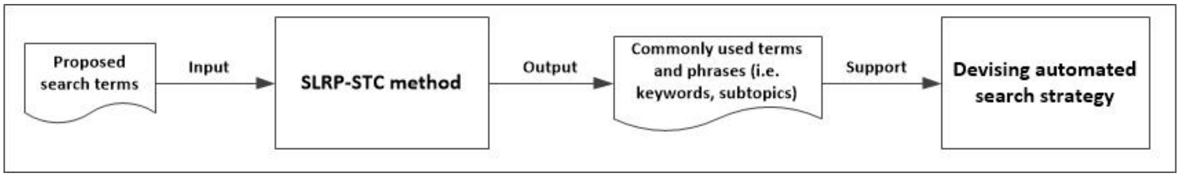

Proposed search terms
Input
SLRP-STC method
Output
Commonly used terms and phrases (i.e. keywords, subtopics)
Support
Devising automated search strategy

Fig. 5: Supporting knowledge discovery process to devise search strategy in Stage 3

3.2 STC Method

This section describes each step in STC method and how the method was applied to support the planning of SLRs. The use of SLRP-STC method requires tool support. A supporting tool named Systematic Literature Review Planning Supporting System (SLRPSS) was developed to implement the SLRP-STC method. SLRPSS is a web-based supporting tool that offers a framework to control various decision-making activities involved in SLR planning phase. It also helps to discover and analyze the necessary information to support SLR planning activities. In order to achieve these functionalities, an open source search results clustering engine, named Carrot^2 [35], was integrated with our supporting tool. Carrot^2 provides functionality to automatically cluster document collections into thematic categories. In addition, a JavaScript library called D3.js [36] was used to enable the information visualization.

Fig. 6 describes the workflow of SLRPSS which include the following main features: 1) it extracts relevant studies and displays them in organized manner; 2) it automatically groups documents into thematic clusters and identifies subtopics of retrieved documents; 3) it displays commonly used terms or phrases and spawns new ideas from visualization of search results; 4) it manages the decision-making activities involved in SLR planning phase.

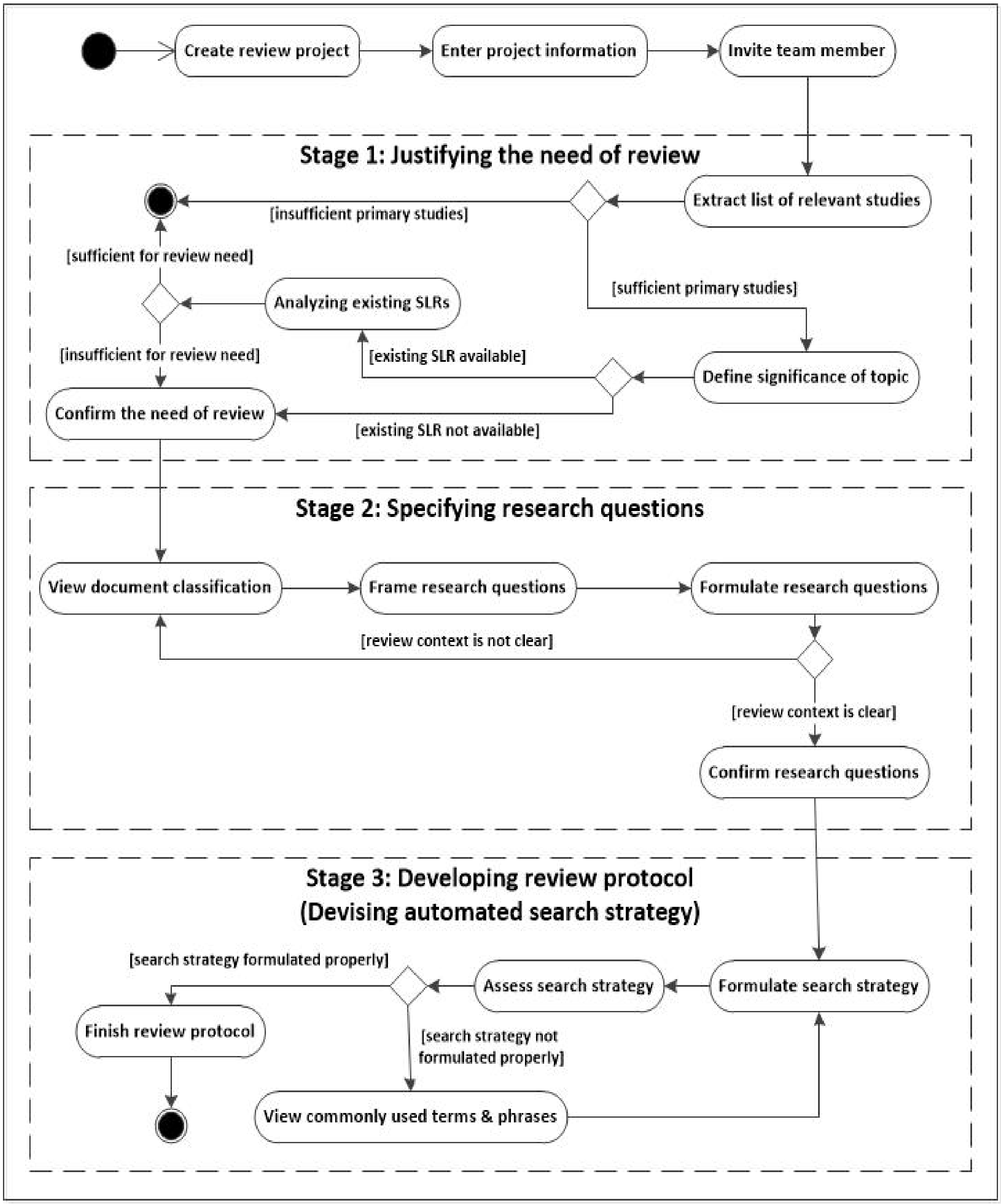

Create review project
Enter project information
Invite team member
Stage 1: Justifying the need of review
Extract list of relevant studies
[insufficient primary studies]
[sufficient for review need]
Analyzing existing SLRs
[sufficient primary studies]
[insufficient for review need]
[existing SLR available]
Define significance of topic
Confirm the need of review
[existing SLR not available]
Stage 2: Specifying research questions
View document classification
Frame research questions
Formulate research questions
[review context is not clear]
[review context is clear]
Confirm research questions
Stage 3: Developing review protocol
(Devising automated search strategy)
[search strategy formulated properly]
Assess search strategy
Formulate search strategy
Finish review protocol
[search strategy not
formulated properly]
View commonly used terms & phrases

Fig. 6: Workflow of SLRPSS

3.2.1 Step 1: Retrieve Web Documents

The STC method starts by providing access to federated search engine that retrieves documents from different data sources. The retrieved documents are added into a document collection in an ordered manner. We consulted SE researchers to provide a list of important digital libraries in the field of SE domain, and reached to a consensus that the five main digital libraries for SE researchers are: IEEE Xplore, ACM Digital Library, Web of Science, Science Direct, Scopus, Google Scholar. Note that, online access to full text of documents in those database requires a valid copyrights and license. Due to the limited access to the full text of documents, we only crawl and scrap citation information, including title, keywords and abstract of documents. The SLRPSS provides a unified search engine wrapper to these digital libraries.

3.2.2 Step 2: Clean Documents

This step starts from the document collection and each document is converted into an ordered string of words. Document processing includes the following sub-steps [27]:

1. **Tokenization:** Sentences are divided into tokens.
2. **Stop-word removal:** It also known as non-word tokenization. There are many words in the document that are meaningless or useless in building suffix tree structure. Numbers, HTML tags, most punctuations and words occur in stop-list (Such as ``are", ``no", ``but", ``however") are stripped.
3. **Stemming:** This step is also called dimensionality reduction. Text in document is replaced with their respective stem. In order to reduce the processing cost without sacrificing processing performance, the system uses light Porter algorithm [31] which is an adapted Porter stemming algorithms [32].

3.2.3 Step 3: Construct Suffix Tree

This step aims to find all base clusters of every document sentences using a suffix tree. To do this, the first thing is to construct a suffix tree by applying the Ukkonen's algorithm [33]. The leaves of tree are marked with a unique sentence identifier that also indicates the corresponding documents. Having built the suffix tree, each internal node of the tree is identified as a base cluster. Second, each internal nodes is assigned with a score that represent the document's usefulness for clustering task. A base cluster will become a maximal phrase cluster if and only if it exceeds a minimal base cluster score. The score function $s(m)$ of base cluster m is derived by the formula [34] as:

$$s(m) = |m| \cdot f(|m_p|) \cdot tfi\,df(w_i) \tag{1}$$

where $s(m)$ is the score of candidate m; $|m|$ is the number of not stop-words phrase terms; $f(|m_p|)$ is the phrase length adjustment; $tfi\,df(w_i)$ is the term frequency adjustment.

3.2.4 Step 4: Merge phrase cluster

Documents are likely to share more than one common phrase. To avoid document overlapping and identical cluster, the fourth step of the STC method is to merge identical base clusters. Only the top scoring clusters will be merged, which is to prevent analysis of studies being influenced by low scoring and less informative clusters. Zamir and Etzioni [27, 28] defined a binary similarity measure between phrase clusters to conduct merging task as:

$$similarity\ (m_i, m_j) = \begin{cases} 1. & if(|m_i \cap m_j| / |m_i| > \alpha) \wedge (|m_i \cap m_j| / |m_j| > \alpha) \\ 0, & otherwise \end{cases} \tag{2}$$

where m_i and m_j are base clusters; $|m_i|$ and $|m_j|$ are the numbers of documents in phrase clusters m_i and m_j respectively; $|m_i \cap m_j|$ is the number of documents common to m_i and m_j; α is the merge threshold. Two base clusters will be merged if and only if they have a similarity of 1. Each merged cluster contains the union of the documents of all its phrase clusters.

3.2.5 Step 5: Study Visualization

In the final step, the system utilizes reviewers strong visual ability to support the knowledge discovery process involved in SLR planning phase. It creates visualization based on phrase clusters to display commonly used terms and phrases that are shared by documents. The document visualization is implemented by using a JavaScript library, named D3.js [36], to combine powerful visualization components and data-driven approach to DOM manipulation. Two common types of visualization are foam tree visualization and circle visualization. Foam tree visualization uses a foam tree to present identified keywords and subtopics, and helps reviewers to efficiently identify the poorly formulated search strings. It is also useful in synonym and abbreviation recognition. On the other hand, circle visualization uses a pie chart to discover commonly used terms, where each sectors of the pie chart represents the identified terms and phrases. The size of each central angle is proportional to the frequency of use of the corresponding terms and phrases.

3.3 An Illustrative Example

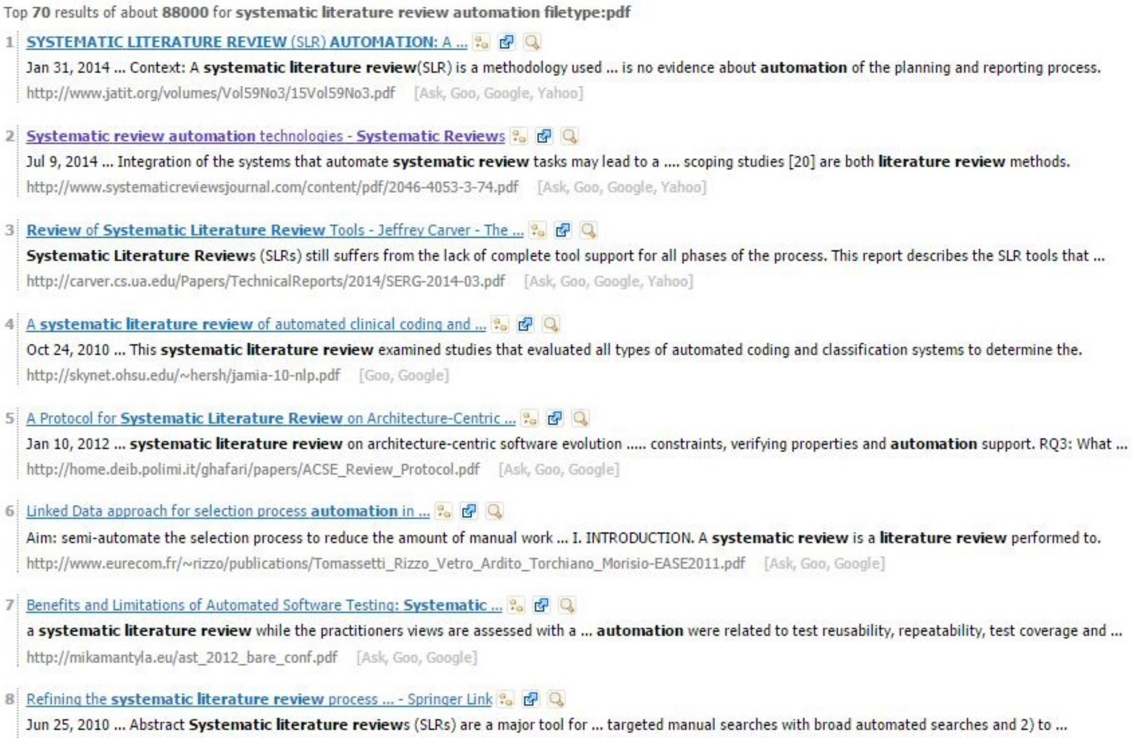
This section illustrates an example of applying SLRP-STC method to plan a SLR. A research topic was selected for this example. The selected topic aims to review the existing automation technologies for supporting SLRs. Through the example, we demonstrated how the SLRP-STC method can be used to plan this SLR in various aspects.

Firstly, general information such as title, project manager, review background, and project timetable were defined to create a new SLR project. An invitation was sent to other system users to conduct a team review on this project. After creating a new project, we started the first stage of our SLR planning with justifying the need of review by checking whether there was sufficient number of primary studies that makes the review worthwhile. In doing this, we sent a query “systematic literature review automation” to seek for potential relevant studies. The SLRPSS tool provides a search engine wrapper to obtain clustering results from Carrot[2]. There were 70 potentially relevant studies retrieved (see Fig. 7). Among them, at least three review papers were identified. The tool also supports to control decision-making activities involved in SLR planning phase. By further analyzing these review papers, we had a basic understanding of the research in the proposed area and we found these review papers mainly focus on research in the field of clinical medicine. As a result, we conclude that there are sufficient primary studies that can be used for conducting a SLR. We found three secondary studies that review current approaches to automate specific SLR activities. Although there are three existing review papers, these papers are associated with medical domain. It is still necessary to conduct a SLR on “SLR automation technology” in the SE domain.

The second stage of planning is to specify research question. This step appears to be sequential, but it is critical to recognize that many activities in this stage involve iteration. During this step, we observed that the retrieved documents were classified into 10 clusters, as shown in Fig. 8. The classification of topics was used to specify reserch questions by identifying review context and narrowing down the review scope.

After specifying the research questions, we are required to define all protocol elements, including search strategy, selection criteria, quality assessment criteria, data extraction, and data synthesis strategy. The phrase clusters show the commonly used terms and phrases shared by these papers and these phrases were potential keywords or subtopics for this SLR. The visualization schema provides reviewers with a strong visual processing abilities to support the knowledge discovery process in SLR planning phase.

The circle view and foam tree view (shown in Fig. 9) provide us a list of identified keywords and subtopics for retrieved documents. We can also use these views to help us devise and revise our search strings in the stage of developing review protocol. In this case, we found many keywords, including “machine learning”, “text mining”, and “visualization” that can be used as search strings because they are frequently used in the research papers. We can also use these keywords and subtopics to evaluate the suitability of the devised search strings. For example, when we used the search string: “systematic literature review automation technology” to search documents, the identified keywords and subtopics for retrieved documents were ‘system’, ‘health’, ‘support’. They are unlikely to describe the subtopics for retrieved documents. Therefore, we can conclude that the proposed search string was poorly formulated (either too broad or inappropriate) and need to be refined.



Top **70** results of about **88000** for **systematic literature review automation filetype:pdf**

1 SYSTEMATIC LITERATURE REVIEW (SLR) AUTOMATION: A ...
Jan 31, 2014 ... Context: A **systematic literature review**(SLR) is a methodology used ... is no evidence about **automation** of the planning and reporting process.
http://www.jatit.org/volumes/Vol59No3/15Vol59No3.pdf [Ask, Goo, Google, Yahoo]

2 Systematic review automation technologies - Systematic Reviews
Jul 9, 2014 ... Integration of the systems that automate **systematic review** tasks may lead to a .... scoping studies [20] are both **literature review** methods.
http://www.systematicreviewsjournal.com/content/pdf/2046-4053-3-74.pdf [Ask, Goo, Google, Yahoo]

3 Review of Systematic Literature Review Tools - Jeffrey Carver - The ...
Systematic Literature Reviews (SLRs) still suffers from the lack of complete tool support for all phases of the process. This report describes the SLR tools that ...
http://carver.cs.ua.edu/Papers/TechnicalReports/2014/SERG-2014-03.pdf [Ask, Goo, Google, Yahoo]

4 A systematic literature review of automated clinical coding and ...
Oct 24, 2010 ... This **systematic literature review** examined studies that evaluated all types of automated coding and classification systems to determine the.
http://skynet.ohsu.edu/~hersh/jamia-10-nlp.pdf [Goo, Google]

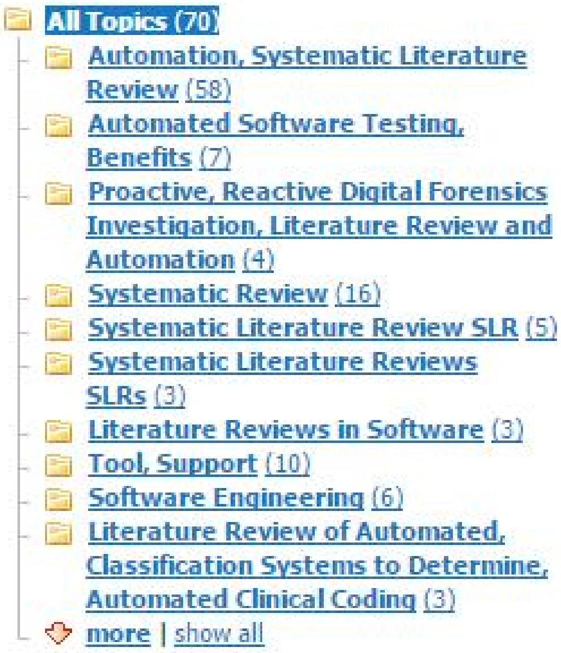
5 A Protocol for Systematic Literature Review on Architecture-Centric ...
Jan 10, 2012 ... **systematic literature review** on architecture-centric software evolution ..... constraints, verifying properties and **automation** support. RQ3: What ...
http://home.deib.polimi.it/ghafari/papers/ACSE_Review_Protocol.pdf [Ask, Goo, Google]

6 Linked Data approach for selection process automation in ...
Aim: semi-automate the selection process to reduce the amount of manual work ... I. INTRODUCTION. A **systematic review** is a **literature review** performed to.
http://www.eurecom.fr/~rizzo/publications/Tomassetti_Rizzo_Vetro_Ardito_Torchiano_Morisio-EASE2011.pdf [Ask, Goo, Google]

7 Benefits and Limitations of Automated Software Testing: Systematic ...
a **systematic literature review** while the practitioners views are assessed with a ... **automation** were related to test reusability, repeatability, test coverage and ...
http://mikamantyla.eu/ast_2012_bare_conf.pdf [Ask, Goo, Google]

8 Refining the systematic literature review process ... - Springer Link
Jun 25, 2010 ... Abstract **Systematic literature reviews** (SLRs) are a major tool for ... targeted manual searches with broad automated searches and 2) to ...

Fig. 7: Extracted list of relevant studies



All Topics (70)
- Automation, Systematic Literature Review (58)
- Automated Software Testing, Benefits (7)
- Proactive, Reactive Digital Forensics Investigation, Literature Review and Automation (4)
- Systematic Review (16)
- Systematic Literature Review SLR (5)
- Systematic Literature Reviews SLRs (3)
- Literature Reviews in Software (3)
- Tool, Support (10)
- Software Engineering (6)
- Literature Review of Automated, Classification Systems to Determine, Automated Clinical Coding (3)

Fig. 8: Classification of relevant studies

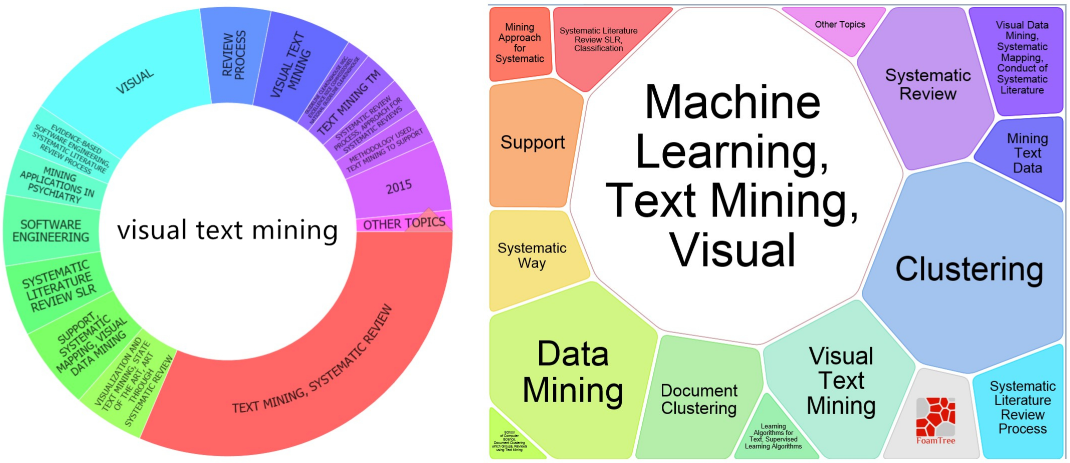


Fig. 9: Identified keywords and subtopics for the example of SLR project

4.0 CASE STUDY

To evaluate the applicability of SLRP-STC method, we conducted a case study with four participants who are postgraduate students with prior experience in performing SLRs. Section 4.1 presents objectives of this case study. The objectives are outlined based on GQM (Goal, Question, Metric) paradigm. Section 4.2 describes preparation and detailed data collection procedures. The collected data are analyzed in Section 4.3.

4.1 Objectives

The main objective of this case study is to evaluate the applicability of SLRP-STC method to support SLR planning activities. Specifically, the goals of case study are defined as follows:

- Goal 1: To validate the applicability of SLRP-STC method in supporting SLR planning process.
- Goal 2: To assess participants’ perceptions of SLRP-STC.

Based on the goals, we defined three questions (Q1, Q2 and Q3) for this case study:

- Q1: Does SLRP-STC method improve efficiency of performing SLR planning process? (Goal 1)
- Q2: Does SLRP-STC method improve quality of SLR planning results? (Goal 1)
- Q3: What are the participants\textquotesingle general perceptions of SLPR-STC method? (Goal 2)

Table 2: Metrics defined for Q1 and Q2

Question	Metric	Description
Q1	M1	The total time taken by Group A to plan the SLR.
	M2	The total time taken by Group B to plan the SLR.
	M3	The average score of the quality of SLRP planning result for Group A.
	M4	The average score of the quality of SLRP planning result for Group B.
Q2	M5	The average score of user acceptance for SLRP-STC method.

In this case study, participants were divided into two groups (Group A and Group B). Group A was required to manually perform the SLR planning without using SLRP-STC method, while Group B was required to perform SLR planning using SLRP-STC method. Table 2 presents the metrics used in this case study to measure the applicability of SLRP-STC method. Table 3 explains the interpretation of comparing the values of two metrics.

Note that β is a constant value that represents the threshold for user acceptance. In this case study, we set $\beta = 3$ because it indicates 'neutral' attitude in the survey.

Table 3: Interpretation of metrics comparisons

Metric Comparison	Interpretation
$M1 \leq M2$	SLRP-STC is unlikely to improve the efficiency of performing SLR planning activities.
$M1 > M2$	SLRP-STC is likely to improve the efficiency of performing SLR planning activities.
$M3 \geq M4$	SLRP-STC is unlikely to improve the quality of undertaking SLR planning activities.
$M3 < M4$	SLRP-STC is likely to improve the quality of undertaking SLR planning activities.
$M5 > \beta$	SLRP-STC method is considered to be applicable.
$M5 \leq \beta$	SLRP-STC method is considered to be inapplicable.

4.2 Data Collection Procedures

The participants selected for this case study were two Ph.D. and two Master's students from the Department of Software Engineering, Faculty of Computer Science and Information Technology (FCSIT), University of Malaya (UM). These postgraduate students were carefully selected by considering their background, knowledge in the field of software engineering (SE), and prior experiences in conducting SLRs and writing SLR research papers. Four participants were divided into two groups: Group A and Group B. Participants from Group A were required to perform SLR planning manually without using the SLRP-STC method, whereas participants from Group B were required to perform SLR planning using the SLRP-STC method. Table 4 presents the background and SLR experience of participants from Group A and Group B. The case study was organized in three phases: 1) Training 2) Execution and 3) Evaluation.

Table 4: List of participants for Group A and Group B

Group No.	Particiapant	Background	Prior Experience in Publishing SLR papers
A (Manual Approach)	Respondent 1	PhD student	Published 2 papers
	Respondent 2	Master's student	Published 1 paper
B (SLRP-STC)	Respondent 3	PhD student	Published 1 paper
	Respondent 4	Master's student	Published 1 paper

4.2.1 Training Phase

During the training phase, all participants were invited to attend a workshop on SLRP-STC method. In order to compare general perceptions of SLRP-STC method between participants from Group A and participants from Group B, it is important to ensure that all participants know exactly what SLRP-STC method is and how to apply it in SLR planning activities. In doing so, the workshop started with a brief introduction of the SLRP-STC method and the case study. In the first session of workshop, a training guide was distributed to help all participants to learn the details of SLRP-STC method. Moreover, an exemplar was used to help participants understand how to apply SLR-STC method in real-world SLR planning projects. In the second session of workshop, participants from Group B were required to continue studying the supporting tool while participants of Group A were allowed to take a short break. A user manual was prepared to serve as a guide in assisting Group B's participants in using the supporting tool for performing the SLR planning activities. During the training process, participants were encouraged to express their opinions and doubts in a free manner.

At the end of the training phase, a questionnaire survey was used to collect the perception of the proposed method from these four participants. The questionnaire survey used a Likert scale for ratings (5-Strongly Agree; 4-Agree; 3-Neutral; 2-Disagree; 1-Strongly Disagree). The questionnaire consists of two open-ended questions that allow participants to express their additional comments and suggestions that will be used to improve the proposed method. The structure of the questionnaire is as follows:

- Section 1 gathers personal information, including background, case study group, prior experience in publishing SLRs.
- Section 2 assesses how participants accept SLRP-STC method. This section consists of 12 questions that helped us to collect general perception of the proposed method from the participants.
- Section 3 consists of two open-ended questions that allow participants to expresses additional comments and suggestions for improving SLRP-STC method.

4.2.2 Execution Phase

During the execution phase, both groups were asked to plan a SLR project on the topic “cloud testing techniques for mobile applications”. The SLR planning template was provided to assist the participants to prepare a SLR planning report that include the SLR planning information and the justifications of the decisions made for their SLR projects during three planning stages. The participants were required to start planning the SLR simultaneously and to finish the SLR planning task in not more than two hours. The time taken for completing the SLR planning project by two groups was recorded. The planning results of two groups were also collected and organized for further evaluation.

4.2.3 Evaluation Phase

In order to evaluate the applicability of SLRP-STC method, data (i.e. SLR planning time) recorded during the execution phase are used for analyzing the efficiency of planning the SLR projects. Two SE researchers who are the experts in conducting SLRs were invited to evaluate the planning reports created by the two groups of participants. The evaluation form was distributed to the SLR experts to collect data to measure the quality of the SLR planning report. The likert scale (5-Strongly Agree; 4-Agree; 3-Neutral; 2-Disagree; 1-Strongly Disagree) was used in the evaluation form for ratings.

The following six questions were asked in the evaluation form to allow the experts to assess the quality of the SLR protocols prepared by the two groups of participants:

1. Q1. The reason for undertaking the review is clear.
2. Q2. The review scope is well defined (neither too broad nor inappropriate).
3. Q3. The review questions are appropriate and clearly formulated.
4. Q4. The automated search strategy used appropriate synonyms, acronyms.
5. Q5. The automated search strategy can be adapted for different databases.
6. Q6. The automated search strings are appropriately formulated (strike a balance between striving for comprehensiveness and maintaining relevance).

4.3 Data Analysis and Results

The results collected during the three phases (i.e. Training, Execution and Evaluation) were analyzed based on the predefined evaluation metrics for the case study, including time taken for planning process to show the efficiency of planning SLR projects, quality of the planning results provided by the SLR experts, and the values of user acceptance for the proposed method provided by the participants of case study through questionnaires.

4.3.1 General Perception of SLRP-STC Method

Four participants (P1, P2, P3 and P4) were required to indicate the level of agreement for general perception of the SLRP-STC method. The questionnaire consists of two parts, first part evaluates the user’s acceptance of

SLRP-STC method, the second part records the suggestions for improving the method. Table 5 presents the ratings provided by the case study participants on the general perception of the proposed method. The survey used Likert-like scale of one to five to represents the level of agreement of participants' perception on the applicability of SLRP-STC method. The table shows the result of participants' level of agreement on the twelve questions related to the perception of the SLRP-STC method. The average score of user acceptance results is 3.7. According to interpretation of metrics comparisons (see Table 3), SLRP-STC method is considered to be applicable because M5 > β.

Table 5: Rating scores for general perception of SLRP-STC method

Question	P1	P2	P3	P4	Mean	Standard Deviation
Q1. This method is easy to understand.	3	4	2	3	3.00	0.816
Q2. This method is useful for supporting SLR planning activities.	4	5	3	4	4.00	0.816
Q3. This method improves the efficiency of performing SLR planning process.	2	3	4	4	3.25	0.957
Q4. This method improves the accuracy of SLR planning results.	4	5	3	4	4.00	0.816
Q5. This method supports effective knowledge acquisition and allows me to quickly extract useful information of existing studies.	4	3	5	4	4.00	0.816
Q6. This method supports effective knowledge discovery process and allows interactive data exploration of existing studies.	3	3	3	4	3.25	0.500
Q7. This method can help me to better understand extracted information.	4	4	5	4	4.25	0.500
Q8. This extracted information supports me to easily justify the need of review.	2	4	3	3	3.00	0.816
Q9. This extracted information supports me to frame/ scope the research questions.	3	4	5	4	4.00	0.816
Q10. This extracted information supports me to devise/refine automated search strings in the stage of protocol development.	4	4	4	5	4.25	0.500
Q11. This method is applicable to the majority of the SLRs in SE domain.	3	3	4	4	3.50	0.577
Q12. This method is flexible to the adapt to different goals and methods between different SLRs.	5	3	5	3	4.00	1.155

4.3.2 Time Taken for Planning An SLR

To determine whether the use of SLRP-STC method can improve the efficiency of the SLR planning process time taken to perform each stage of SLR planning is measured. The result is tabulated in Table 6. The total time spent on SLR planning process is 87 minutes (M1) for group A and 75 minutes (M2) for group B. Based on metrics comparisons in Table 3, it is conclusive to say that SLRP-STC is likely to improve the efficiency of performing SLR planning activities because M1 > M2.

4.3.3 Quality of SLR Planning Reports

To determine whether the quality of planning results can be improved by using SLRP-STC method, the measurement collected from two SLR experts (E1 and E2) using an evaluation survey form. The survey used Likert scale from one to five to represents the level of agreement of experts on the quality of planning results. The results are shown in Table 7. In general, the average score for Group A (Manual approach) is 17.5 (M3) and for Group B (SLRP-STC) is 25 (M4). Based on the interpretation of metrics comparisons (see Table 3), , it is conclusive to say that SLRP-STC is likely to improve the quality of undertaking SLR planning activities since M3 < M4.

Table 6: Time taken to perform each stage of SLR planning (measured in minutes)

SLR Planning Stage	**Group A** **(Manual Approach)**	**Group B** **(SLRP-STC)**
Stage 1: Justifying the need of review	63	41
Stage 2: Specifying research questions	9	25
Stage 3: Devising automated search strategy	15	9
Total time taken	**87**	**75**

Table 7: Rating scores to assess quality of SLR planning protocols by expert 1 and Expert 2

Evaluation Question	**Protocol 1 (Group A)**		**Mean Score (Protocol 1)**	**Protocol 2 (Group B)**		**Mean Score (Protocol 2)**
	E1	**E2**		**E1**	**E2**	
Q1. The reason for undertaking the review is clear.	3	3	3	4	5	4.5
Q2. The review scope is well defined (neither too broad nor inappropriate)	3	2	2.5	4	4	4
Q3. The review questions are appropriate and clearly formulated.	3	4	3.5	4	4	4
Q4. The automated search strategy used appropriate synonyms, acronyms.	3	4	3.5	4	5	4.5
Q5. The automated search strategy adapted as needed for appropriate databases.	2	3	2.5	4	5	4.5
Q6. The automated search strings are appropriated formulated (strike a balance between striving for comprehensiveness and maintaining relevance).	2	3	2.5	3	4	3.5
Total Score	**16**	**19**	**17.5**	**23**	**27**	**25**

5.0 DISCUSSION

This case study aims to evaluate the applicability of the SLRP-STC method from three aspects, including the time taken for planning process, quality of planning results, and participants' general perception on the proposed method. The results demonstrate that it is applicable for research students to apply the proposed method in SLR planning phase. This section discusses the case study results and how the results answer the three questions defined in Section 4.1.

5.1 Does SLRP-STC method improve efficiency of performing SLR planning process?

Compared with the total time spent for planning process for Group A, Group B spent less time in SLR planning process, especially at stage 1 and stage 3. This shows that SLRP-STC method is likely to save time and efforts in justifying the need of review and devising automated search strategy. Group B took longer time in stage 2 of SLR planning phase specifying the research question. However, the quality of research question of Group B was better than Group A. The reason of using SLRP-STC method might takes longer time to specify research question is that SLRP-STC method helps reviewers to consider the research questions from four PICO aspects, namely, population, intervention, context, and outcome. Such practice is useful and necessary in helping reviewers to frame answerable research questions. On the one hand, it is of great importance to formulate the appropriate research questions during the SLR process because research questions are key element of the entire SLR methodology. On the other hand, well-defined research questions helps to reduce the workload of performing the overall SLR process. This prevents the reviewers from keep changing the research questions during SLR.

5.2 Does SLRP-STC method improve quality of SLR planning results?

The average score of Group A (manual) is less than Group B (SLRP-STC). More specifically, in Question 1, the mean score for Group A is higher than Group B. We can speculate that experts agreed that Group B, after applying SLRP-STC method, presented a clearer reason for undertaking the review. Likewise, in Question 2, we can learn that experts agreed the Group B has better defined research scope. For Question 3, the higher score of Group B suggests that SLR experts believed that Group B had produced better research questions. In Question 4 and Question 5, experts agreed that the quality of automated search strategy for Group B was better than Group A. Hence, from the evaluation results, SLRP-STC method is likely to improve the quality of SLR planning results.

5.3 What are the participants' general perceptions of SLPR-STC method?

The results show that participants agreed that SLRP-STC method was useful for supporting SLR planning activities. Additionally, the participants agree that SLRP-STC method can improve the efficiency of performing SLR planning process and improve the accuracy of the SLR planning. Furthermore, the participants agreed that SLRP-STC method supported effective knowledge acquisition process, especially in extracting information that can support to devise or refine automated search strings in the stage of protocol development. Also, the participants agreed that SLRP-STC method was applicable to planning SLRs in the SE domain. SLRP-STC method was flexible to be adapted to different goals and methods between different SLRs. As the average score of the participants is greater than β value, SLRP-STC method is applicable to support the planning of SLRs.

There are a few recommendations suggested by participants. Their suggestions will be taken into consideration to improve and refine SLRP-STC method. One of the participants suggested that we should focus on support planning decision-making activities by defining a number of reasonable criteria. Another participant believed that it would be more interesting if additional text mining features can be added to support the SLR planning activities.

5.4 Theoretical contributions

Prior empirical software engineering (ESE) research had not explained enough how text mining techniques can fit into the planning of an SLR [18-25]. TM techniques enable effective knowledge acquisition and discovery process to support SLRs and reduce the workload in planning SLRs. This research provides an improved understanding about the relationship between SLR planning and quality of SLR through an informed decision-making perspective based on TM techniques. The proposed method addresses the requirements of Web document clustering to present important information that contributes to the quality improvement of SLR review protocols. The research provides a new way of thinking about how the level of quality of a review can be improved, by reducing the decision-making problems that can arise during SLR planning, especially in the development of review protocols. This is because, the existing studies [18-25] have not considered applying TM tecniques to retrieve information that can support the decision-making activities during SLR planning. Although Fernández-Sáez et al. [20] have developed a tool named SLR-Tool to capture the planning information of a project, they did not TM technique to support effective planning. Thus, the STC method applied in clustering web documents provides a new basis for understanding how automatic document clustering fit into the traditional SLR planning process. Additionally, this research explores how the visualization of document clustering information may improve the quality of a review by defining the correct review protocol elements, including search strategy, selection criteria, quality assessment criteria, data extraction, and data synthesis strategy during SLR planning stage. The SLRP-STC Method provides a basis to support future research in SLR planning, which warrants further investigation in developing a good review protocol during SLR planning, rather on a total focus of ESE research in conducting and reporting the SLR.

6.0 CONCLUSION AND FUTURE WORK

SLR is a research methodology used to review the relevant research studies and synthesize the best evidence to address specific research questions, and identify research gaps. It is primarily an information-based process, where each stage must be fairly and systematically designed in planning phase. Due to the comprehensive and rigorous process of conducting a SLR, exhaustive planning activities are required. Applying effective TM techniques is the best solution to facilitate the decision-making activities involved during the SLR planning phase.

In this research, we proposed SLRP-STC method to support the effective knowledge acquisition and knowledge discovery process for large Web document collections. This method aims to support the researchers and students to make informed decision during the SLR planning phase. Next, a web-based supporting tool was developed to implement the proposed method. At the end of this research, a case study was conducted to evaluate the proposed method. The case study was conducted to measure the applicability aspects of the SLRP-STC method, including the time taken for SLR planning phase, the quality of planning results, between SLRP-STC method manual method, and the participants' general perception on SLRP-STC method. The experimental results show that the SLRP-STC method is useful in improving the performance of planning an SLR. Also, the participants' feedback shows that it is applicable to use SLRP-STC method during SLR planning phase to make informed decision-making in planning a SLR.

This research contributes to encourage the adoption of SLR methodology by facilitating the SLR planning process. In addition, this research applies a text mining method, STC that supports effective knowledge acquisition and knowledge discovery process involved in SLR planning phase. The proposed method, SLRP-STC contributes to support the data exploratory and data analysis involved during the SLR planning phase by clustering and mining the Web retrieved documents automatically. This method contributes to help researchers and students in three ways: 1) to extract research topics easily; 2) to better interpret extracted information; and 3) to quickly realize the proposed search term is poorly formulated or inappropriate, and refine the search strings. The use of SLRP-STC method can significantly improves the performance of planning SLRs. There are some limitations of the SLRP-STC method in some aspects. Firstly, the SLRP-STC method can only be used to partially support the SLR planning phase. Only the first three stages are facilitated by the proposed method. The fourth stage of planning phase is not considered in the scope of this research. In addition, the size of the participants set in the case studies is small. There were only two participants that act as reviewers for each group but the results show that is it applicable to more general case.

For future work, there are many ways to enhance the proposed work. The possible enhancements for this research includes; 1) extending the research to cover the final stage of SLR planning phase and other aspects of developing review protocol, including defining selection strategy, data extraction and synthesis strategies; 2) extending the research to apply STC method to support other phases of the SLR process, in particular, to support search and selection activities; 3) extending the research to integrate our work with other TM techniques or tools for supporting SLR planning phase; 4) extending the research to add more visualization features, such as citation map and citation network; 5) comparing the proposed method with other text mining techniques in addition to manual method; and 6) extending the proposed method by integrating the method with other text mining techniques. It is expected that many useful and novel visualization techniques will emerge over times as researchers continue exploring the text mining applications.

ACKNOWLEDGEMENTS

This work is supported by a University of Malaya Research Grant (UMRG), Project Code: RP028C-14HTM.

REFERENCES

[1] B. A. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman， "Systematic literature reviews in software engineering-a systematic literature review". *Information and Software Technology*, Vol. 51, No. 1, 2009, pp. 7-15.

[2] M. Staples and M. Niazi, "Experiences using systematic review guidelines". *Journal of Systems and Software*, Vol. 80, No. 9, 2007, pp. 1425-1437.

[3] H. Ramampiaro, D. Cruzes, R. Conradi, and M. Mendona, "Supporting evidence-based software engineering with collaborative information retrieval", in *6th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), Chicago, IL*, 2010, pp. 1-5.

[4] Y. Aphinyanaphongs, A. Statnikov, and C. F. Aliferis, "A comparison of citation metrics to machine learning filters for the identification of high quality medline documents". *Journal of the American Medical Informatics Association*, Vol. 13, No. 4, 2006, pp. 446-455.

[5] B. A. Kitchenham, *Procedures for Performing Systematic Reviews*. Joint Technical Report, Keele University TR/SE-0401 and NICTA 0400011T.1, July 2004.

[6] P. Brereton, B. A. Kitchenham, D. Budgen, M. Turner, and M. Khalil, "Lessons from applying the systematic literature review process within the software engineering domain". *Journal of Systems and Software*, Vol. 80, No. 4, 2007, pp. 571-583.

[7] E. Hassler, J. C. Carver, N. A. Kraft, and D. Hale, "Outcomes of a community workshop to identify and rank barriers to the systematic literature review process", in *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, New York*, ACM, 2014, Article 31, 10 pages.

[8] R. Feldman, and J. Sanger, *The text mining handbook: advanced approaches in analyzing unstructured data*. Cambridge University Press, 2007.

[9] M. W. Berry, "Survey of text mining". *Computing Reviews*, Vol. 45, No. 9, 2004, p. 548.

[10] A. Hotho, A. Nürnberger, and G. Paaß, "A Brief Survey of Text Mining". *LDV Forum - GLDV Journal for Computational Linguistics and Language Technology*, Vol. 20, No. 1, 2005, pp. 19-62.

[11] V. Gupta and G. S. Lehal, "A survey of text mining techniques and applications". *Journal of Emerging Technologies in Web Intelligence*, Vol. 1, No. 1, 2009, pp. 60-76.

[12] A. K. Nassirtoussi, S. Aghabozorgi, T. Y. Wah and D. C. L. Ngo, "Text mining for market prediction: A systematic review". *Expert Systems with Applications*, Vol. 41, No. 16, 2014, pp. 7653-7670.

[13] N. S. Safa, N. A. Ghani, M. A. Ismail, "An artificial neural network classification approach for improving accuracy of customer identification in e-commerce". *Malaysian Journal of Computer Science*, Vol. 27, No. 3, 2014, pp.171-185.

[14] T. Zia, and M. P. Akhter and Q. Abbas, "Comparative Study of Feature Selection Approaches for Urdu Text Categorization". *Malaysian Journal of Computer Science*, Vol. 28, No. 2, 2015, pp. 93-109.

[15] L. Feng, *Suffix Tree Clustering Method to Support Systematic Literature Review Planning Phase (Chapter 3)*. Unpublished Master's Thesis, University of Malaya, Kuala Lumpur, Malaysia, 2015, pp. 28-46.

[16] K.D. Varathan, T. M. T. Sembok, R. A. Kadir and N. Omar, “Semantic indexing for question answering system”. *Malaysian Journal of Computer Science*, Vol. 27, No. 4, 2014, pp. 261-274.

[17] B. A. Kitchenham, D. Budgen, P. Brereton, *Evidence-Based Software Engineering and Systematic Reviews (Vol. 4)*. CRC Press, 2015, pp.39-54

[18] V. Malheiros, E. Hohn, R. Pinho, and M. Mendonca, “A visual text mining approach for systematic reviews”, in *First International Symposium on Empirical Software Engineering and Measurement (ESEM), Madrid*, 2007, pp. 245-254.

[19] K. R. Felizardo, E. Y. Nakagawa, D. Feitosa, R. Minghim, and J. C. Maldonado, “An approach based on visual text mining to support categorization and classification in the systematic mapping”, in *Proceedings of the 14th International Conference on Evaluation and Assessment in Software Engineering (EASE), British Computer Society, Swinton, UK*, 2010, pp. 34-43.

[20] A. M. Fernández-Sáez, M. G. Bocco, F. P. Romero, “SLR-Tool: A Tool for Performing Systematic Literature Reviews”, in *5th International Conference on Software and Data Technologies (ICSOFT), Vol. 2, Athens, Greece*, 2010, pp. 157-166.

[21] F. Tomassetti, G. Rizzo, A. Vetro, L. Ardito, M. Torchiano, and M. Morisio, “Linked data approach for selection process automation in systematic reviews”, in *15th Annual Conference on Evaluation and Assessment in Software Engineering (EASE), Durham*, 2011, pp. 31-35.

[22] M. Ghafari, M. Saleh, and T. Ebrahimi, “A federated search approach to facilitate systematic literature review in software engineering”. *International Journal of Software Engineering and Applications (IJSEA)*, Vol. 3, No. 2, 2012, pp. 13-24.

[23] S. Fabbri, E. Hernandes, A. Di Thommazo, A. Belgamo, A. Zamboni, and C. Silva, “Using information visualization and text mining to facilitate the conduction of systematic literature reviews”, in *14th International Conference on Enterprise Information System (ICEIS), Wroclaw, Poland*, 2012, pp. 243-256.

[24] J. Torres, D. Cruzes, and L. Salvador, “Automatically locating results to support systematic reviews in software engineering”, in *Workshop Latinoamericano Ingeniería de Software Experimental (ESELAW), Montevideo, Uruguay*, 2013, pp, 6-19.

[25] K. R. Felizardo, E. Y. Nakagawa, S. G. MacDonell, and J. C. Maldonado, “A visual analysis approach to update systematic reviews”, in *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE), New York*, ACM, 2014, Article 4, 10 pages.

[26] D. Gusfield, *Algorithms on strings, trees and sequences: computer science and computational biology*. Cambridge University Press, 1997.

[27] O. E. Zamir and O. Etzioni, “Web document clustering: A feasibility demonstration”, in *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in Information Retrieval, New York*, ACM, 1998, pp. 46-54.

[28] O. E. Zamir, *Clustering web documents: a phrase-based method for grouping search engine results*. Doctoral dissertation, University of Washington, 1999, p. 56.

[29] Keele Staffs, *Guidelines for performing systematic literature reviews in software engineering*. Technical report, Ver. 2.3, EBSE Technical Report, 2007.

[30] B. A. Kitchenham and P. Brereton, “A systematic review of systematic review process research in software engineering”. *Information and Software Technology*, Vol. 55, No. 12, 2013, pp. 2049-2075.

[31] L. S. Larkey, L. Ballesteros, and M. E. Connell, “Improving stemming for arabic information retrieval: light stemming and co-occurrence analysis”, in *Proceedings of the 25th Annual International ACM*

SIGIR Conference on Research and development in Information Retrieval, New York, ACM, 2002, pp. 275-282.

[32] M. F. Porter, "An algorithm for suffix stripping". *Program (Automated Library and Information Systems)*, Vol. 14, No. 3, 1980, pp. 130-137.

[33] E. Ukkonen, "Algorithms for approximate string matching". *Information and Control*, Vol. 64, No. 1, 1985, pp. 100-118.

[34] R. C. Pushplata, "Auto-assemblage for Suffix Tree Clustering". *International Journal of Advanced Research in Computer Engineering & Technology*, Vol. 1, No. 4, 2012, pp. 600-608.

[35] Carrot Search, *http://project.carrot2.org/index.html*

[36] M. Bostock, V. Ogievetsky, and J. Heer, "D3 data-driven documents". *IEEE Transactions on Visualization and Computer Graphics*, Vol. 17, No. 12, 2011, pp. 2301-2309.