

# What papers should I cite from my reading list?

## User evaluation of a manuscript preparatory assistive task

Foo, Schubert; Pang, Natalie; Sesagiri Raamkumar, Aravind

2016

Sesagiri Raamkumar, A., Foo, S., & Pang, N. (2016). What papers should I cite from my reading list? User evaluation of a manuscript preparatory assistive task. Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraies, 51–62.

<https://hdl.handle.net/10356/81351>

---

© 2016 The Author(s) (published by ACM). This is the author created version of a work that has been peer reviewed and accepted for publication in Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraies, published by ACM on behalf of the author. It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [<http://ceur-ws.org/Vol-1610/paper7.pdf>].

*Downloaded on 17 Apr 2021 08:25:42 SGT*

# What papers should I cite from my reading list? User evaluation of a manuscript preparatory assistive task

Aravind Sesagiri Raamkumar, Schubert Foo, Natalie Pang

Wee Kim Wee School of Communication and Information,  
Nanyang Technological University, Singapore  
{aravind002,sfoo,nlspang}@ntu.edu.sg

**Abstract.** Literature Review (LR) and Manuscript Preparatory (MP) tasks are two key activities for researchers. While process-based and technological-oriented interventions have been introduced to bridge the apparent gap between novices and experts for LR tasks, there are very few approaches for MP tasks. In this paper, we introduce a novel task of shortlisting important papers from the reading list of researchers, meant for citation in a manuscript. The technique helps in identifying the important and unique papers in the reading list. Based on a user evaluation study conducted with 116 participants, the effectiveness and usefulness of the task is shown using multiple evaluation metrics. Results show that research students prefer this task more than research and academic staff. Qualitative feedback of the participants including the preferred aspects along with critical comments is presented in this paper.

**Keywords:** manuscript preparation; shortlisting citations; scientific paper information retrieval; scientific paper recommender systems; digital libraries

## 1 Introduction

The Scientific Publication Lifecycle comprises of different activities carried out by researchers [5]. Of all these activities, the three main activities are literature review, actual research work and dissemination of results through conferences and journals. These three activities in themselves cover multiple sub-activities that require specific expertise and experience [16]. Prior studies have shown researchers with low experience, face difficulties in completing research related activities [9, 15]. These researchers rely on assistance from supervisors, experts and librarians for learning the required skills to pursue such activities. Scenarios where external assistance have been traditionally required are (i) selection of information sources (academic search engines, databases and citation indices), (ii) formulation of search queries, (iii) browsing of retrieved results and (iv) relevance judgement of retrieved articles [9]. Apart from human assistance, academic assistive systems have been built for alleviating the expertise gap between experts and novices in terms of research execution. Some of these interventions include search systems with faceted user interfaces for better dis-

play of search results [2], bibliometric tools for visualizing citation networks [7] and scientific paper recommender systems [3, 14], to name a few.

In the area of manuscript writing, techniques have been proposed to recommend articles for citation contexts in manuscripts [11]. In the context of manuscript publication, prior studies have tried to recommend prospective conference venues [25] most suited for the research in hand. One unexplored area is helping researchers in identifying the important and unique papers that can be potentially cited in the manuscript. This identification is affected by two factors. The first factor is the type of research where citation of a particular paper makes sense due to the particular citation context. The second factor is the type of article (for e.g., conference full paper, journal paper, demo paper) that the author is intending to write. For the first factor, there have been some previous studies [11, 14, 21]. The second factor represents a task that can be explored since the article-type places a constraint on the citations that can be made in a manuscript, in terms of dimensions such as recency, quantity, to name a few.

In our research, we address this new manuscript preparatory task with the objective of shortlisting papers from the reading list of researchers based on article-type preference. By the term ‘shortlisting’, we allude to the nature of the task in identifying important papers from the reading list. This task is part of a functionality provided by an assistive system called Rec4LRW meant for helping researchers in literature review and manuscript preparation. The system uses a corpus of papers, built from an extract of ACM Digital Library (ACM DL). It is hypothesized that the Rec4LRW system will be highly beneficial to novice researchers such as Ph.D. and Masters students and also for researchers who are venturing into new research topics. A user evaluation study was conducted to evaluate all the tasks in the system, from a researcher’s perspective.

In this paper, we report the findings from the study. The study was conducted with 116 participants comprising of research students, academic staff and research staff. Results from the six evaluation measures show that the participants prefer to have the shortlisting feature included in academic search systems and digital libraries. Subjective feedback from the participants in terms of the preferred features and the features that need to be improved, are also presented in the paper.

The remainder of this work is organized as follows. Section two surveys the related work. The Rec4LRW system is introduced along with dataset, technical details and unique UI features in section three. In section four, the shortlisting technique of the task is explained. Details about the user study and data collection are outlined in Section five. The evaluation results are presented in section six. The concluding remarks and future plans for research are provided in the final section.

## **2 Related Work**

Conceptual models and systems have been proposed in the past for helping researchers during manuscript writing. Generating recommendations for citation contexts is an approach meant to help the researcher in finding candidate citations for particular placeholders (locations) in the manuscript. These studies make use of content oriented recommender techniques as there is no scope for using Collaborative Filtering (CF)

based techniques due to lack of user ratings. Translation models have been specifically used in [13, 17] as they are able to handle the issue of vocabulary mismatch gap between the user query and document content. The efficiency of the approaches is dependent on the comprehensiveness of training set data as the locations and corresponding citations data are recorded. The study in [11] is the most sophisticated, as it does not expect the user to mark the citation contexts in the input paper unlike other studies where the contexts have to be set by the user. The proposed model in the study learns the placeholders in previous research articles where citations are widely made so that the citation recommendation can be made on occurrence of similar patterns. The methods in these studies are heavily reliant on the quality & quantity of training data; therefore they are not applicable to systems which lack access to full text of research papers.

Citation suggestions have also been provided as part of reference management and stand-alone recommendation tools. ActiveCite [21] is a recommendation tool that provides both high level and specific citation suggestions based on text mining techniques. Docear is one of the latest reference management software [3] with a mind map feature that helps users in better organizing their references. The in-built recommendation module in this tool is based on Content based (CB) recommendation technique with all the data stored in a central server. The Refseer system [14], similar to ActiveCite, provides both global and local (particular citation context) level recommendations. The system is based on the non-parametric probabilistic model proposed in [12]. These systems depend on the quality and quantity of full text data available in the central server as scarcity of papers could lead to redundant recommendations.

Even though article-type recommendations have not been practically implemented, the prospective idea has been discussed in few studies. The article-type dimension has been highlighted as part of the user's 'Purpose' in the multi-layer contextual model put forth in [8] and as one of the facets in document contextual information in [6]. The article type indirectly refers to the goal of the researcher. It is to be noted that goal or purpose related dimensions have been considered for research in other research areas of recommender systems namely course recommendations [23] and TV guide recommendations [20]. Our work, on the other hand, is the first to explore this task of providing article-type based recommendations with the aim of shortlisting important and unique papers from the cumulative reading list prepared by researchers during their literature review. Through this study, we hope to open new avenues of research which requires a different kind of mining of bibliographic data, for providing more relevant results.

### **3 Assistive System**

#### **3.1 Brief Overview**

The Rec4LRW system has been built as a tool aimed to help researchers in two main tasks of literature review and one manuscript preparatory task. The three tasks are (i) Building an initial reading list of research papers, (ii) Finding similar papers based on a set of papers, and (iii) Shortlisting papers from the final reading list for inclusion in

manuscript based on article-type choice. The usage context of the system is as follows. Typically, a researcher would run the first task for one or two times at the start of the literature review, followed by selection of few relevant seed papers which are then used for task 2. The second task takes these seed papers as an input to find topically similar papers. This task is run multiple times until the researcher is satisfied with the whole list of papers in the reading list. The third task (described in this paper), is meant to be run when the researcher is at the stage of writing manuscripts for publication. It is observed that the researcher would maintain numerous papers in his/her reading list while performing research (could be more than 100 papers for most research studies). The third task helps the researcher in identifying both important and unique papers from the reading list. The shortlisted papers count varies as per the article-type preference of the researcher. The recommendation mechanisms of the three tasks are based on seven features/criteria that represent the characteristics of the bibliography and its relationship with the parent research paper [19].

### **3.2 Dataset**

A snapshot of the ACM Digital Library (ACM DL) is used as the dataset for the system. Papers from proceedings and journals for the period 1951 to 2011 form the dataset. The papers from the dataset have been shortlisted based on full text and metadata availability in the dataset, to form the sample set/corpus for the system. The sample set contains a total of 103,739 articles and corresponding 2,320,345 references.

### **3.3 User-Interface (UI) Features**

In this sub-section, the unique UI features of the Rec4LRW system are presented. Apart from the regular fields such as author name(s), abstract, publication year and citation count, the system displays the fields:- author-specified keywords, references count and short summary of the paper (if the abstract of the paper is missing). Most importantly, we have included information cue labels beside the title for each article. There are four labels (1) Popular, (2) Recent, (3) High Reach and (4) Survey/Review. A screenshot from the system for the cue labels (adjacent to article title) is provided in Figure 1.

The display logic for the cue labels are described as follows. The recent label is displayed for papers published between the years 2009 and 2011 (the most recent papers in the ACM dataset is of 2011). The survey/review label is displayed for papers which are of the type - literature survey or review. For the popular label, the unique citation counts of all papers for the selected research topic are first retrieved from the database. The label is displayed for a paper if the citation count is in the top 5% percentile of the citation counts for that topic. Similar logic is used for the high reach label with references count data. The high reach label indicates that the paper has more number of references than most other articles for the research topic, thereby facilitating the scope for extended citation chaining. Specifically for task 3, the system provides an option for the user to view the papers in the parent cluster of the

shortlisted papers. This feature helps the user in serendipitously finding more papers for reading. The screenshot for this feature is provided in Figure 1.

**Shortlisted papers based on the article-type preference**

- 1) [Measurement and analysis of online social networks](#) Popular  
 Alan Mislove, Massimiliano Marcon; Krishna P. Gummadi; Peter Druschel; Bobby Bhattacharjee - Internet measurement, 2007  
**Abstract:** Online social networking sites like Orkut, YouTube, and Flickr are among the most popular sites on the Internet. Understanding these graphs is important, both applications of online social networks.  
**Author Specified Keywords:** analysis; measurement; social networks  
**Citation Count:** 115 **References Count:** 55 [View papers in the parent cluster](#)

Information cue labels
- 2) [SybilGuard: defending against sybil attacks via social networks](#) Popular  
 Haifeng Yu; Michael Kaminsky; Phillip B. Gibbons; Abraham Flaxman - SIGCOMM Comput. Commun. Rev., 2006  
**Abstract:** Peer-to-peer and other decentralized, distributed systems are known to be particularly vulnerable to sybil attacks. In a sybil attack, a malicious user obtains mu distinct nodes in the system. By controlling a large fraction of the nodes in the system, the malicious user is able to "out vote" the honest users in collaborative tasks such presents SybilGuard, a novel protocol for limiting the corruptive influences of sybil attacks. Our protocol is based on the "social network" among user identities, where an established trust relationship. Malicious users can create many identities but few trust relationships. Thus, there is a disproportionately-small "cut" in the graph between SybilGuard exploits this property to bound the number of identities a malicious user can create. We show the effectiveness of SybilGuard both analytically and experim  
**Author Specified Keywords:** social networks; sybil attack; sybil identity; sybilGuard  
**Citation Count:** 40 **References Count:** 27 [View papers in the parent cluster](#)

Parent cluster of the shortlisted paper
- 3) [Structure and evolution of online social networks](#) Popular  
 Ravi Kumar; Jasmine Novak; Andrew Tomkins - Knowledge discovery and data mining, 2006  
**Abstract:** In this paper, we consider the evolution of structure within large online social networks. We present a series of measurements of two such networks, together c ten million friendship links, annotated with metadata capturing the time of every event in the life of the network. Our measurements expose a surprising segmentation of who do not participate in the network; isolated communities which overwhelmingly display star structure; and a giant component anchored by a well-connected core reg stars. We present a simple model of network growth which captures these aspects of component structure. The model follows our experimental results, characterizing us inviters who encourage offline friends and acquaintances to migrate online; and linkers who fully participate in the social evolution of the network.  
**Author Specified Keywords:** graph evolution; graph mining; small-world phenomenon; social networks; stars  
**Citation Count:** 77 **References Count:** 30 [View papers in the parent cluster](#)

Fig. 1. Sample list of shortlisted papers for the task output

#### 4 Technique For Shortlisting Papers From Reading List

The objective of this task is to help researchers in identifying important (based on citation counts) and unique papers from the final reading list. These papers are to be considered as potential candidates for citation in the manuscript. For this task, the Girvan–Newman algorithm [10] was used for identifying the clusters in the citations network. The specific goal of clustering is to identify the communities within the citation network. From the identified clusters, the top cited papers are shortlisted. The algorithm is implemented as the *EdgeBetweennessClusterer* in JUNG library. The algorithm was selected as it is the one of the most prominent community detection algorithms based on link removal. The other algorithms considered were voltage clustering algorithm [24] and bi-component DFS clustering algorithm [22]. Based on internal trail tests, the Girvan–Newman algorithm was able to consistently identify meaningful clusters using the graph constructed with the citations and references of the papers from the reading list.

As a part of this task, we have tried to explore the notion of varying the count of shortlisted papers by article-type choice. For this purpose, four article-types were considered: conference full paper (cfp), conference poster (cp), generic research paper

(gp)<sup>1</sup> and case study (cs). The article-type classification is not part of the ACM metadata but it is partly inspired by the article classification used in Emerald publications. The number of papers to be shortlisted for these article-types was identified by using the historical data from ACM dataset. First, the papers in the dataset were filtered by using the title field and section field for the four article-types. Second, the average of the references count was calculated for the filtered papers for each article-type from previous step. The average references count for the article-types gp, cs, cfp and cp are 26, 17, 16 and 6 respectively. This new data field is used to set the number of papers to be retrieved from the paper clusters. The procedure for this technique is given in Procedure 1.

---

**Procedure 1** shortlistpapers( $P$ )

---

Input:  $P$  – set of papers in the final reading list  
 $AT$  – article-type choice of the user

- 1:  $RC \leftarrow$  the average references count retrieved for  $AT$
- 2:  $R \leftarrow$  list of retrieved citations & references of papers from  $P$
- 3:  $G \leftarrow$  directed sparse graph created with papers from  $R$
- 4: run *edge betweenness* algorithm on  $G$  to form cluster set  $C$
- 5:  $S \leftarrow$  final list of shortlisted papers
- 6: **if**  $|C| > RC$  **then**
- 7: **while**  $|S| = RC$
- 8: **for** each cluster in  $C$  **do**
- 9: sort papers in the cluster on citation count
- 10:  $s \leftarrow$  top ranked paper from the cluster
- 11: add  $s$  to  $S$
- 12: **end for**
- 13: **end while**
- 14: **else**
- 15:  $N \leftarrow 0$
- 16: **while**  $|S| = RC$
- 17:  $N \leftarrow N + 1$
- 18: **for** each cluster in  $C$  **do**
- 19: sort papers in the cluster on citation count
- 20:  $s \leftarrow N$  ranked paper from the cluster
- 21: add  $s$  to  $S$
- 22: **end for**
- 23: **end while**
- 24: **end if**
- 25: display papers from  $S$  to user

---



---

<sup>1</sup> A paper is qualified as a generic research paper when it doesn't fall quality under the requirements of all the other article-types

## 5 User Evaluation Study

In IR and RS studies, offline experiments are conducted for evaluating the proposed technique/algorithm with baseline approaches. Since the task addressed in the current study is a novel task, the best option was to perform a user evaluation study with researchers. Considering the suggestions from [4], the objective of the study was to ascertain the usefulness and effectiveness of the task to researchers. The specific evaluation goals were (i) ascertain the agreement percentages of the evaluation measures and (ii) identify the top preferred and critical aspects of the task through the subjective feedback of the participants. An online pre-screening survey was conducted to identify the potential participants. Participants needed to have experience in writing conference or journal paper(s) as a qualification for taking part in the study.

All the participants were required to evaluate the three tasks and the overall system. In task 1, the participants had to select a research topic from a list of 43 research topics. On selection of topic, the system provides the top 20 paper recommendations which are meant to be part of the initial LR reading list. In task 2, they had to select a minimum of five papers from task 1 in order for the system to retrieve 30 topically similar papers. For the third task, the participants were requested to add at least 30 papers in the reading list. The paper count was set to 30 as the threshold for highest number of shortlisted papers was 26 (for the article-type ‘generic research paper’). The three other article-types provided for the experiment were conference full paper, conference poster and case study. The shortlisted papers count for these article-types was fixed by taking average of the references count of the related papers from the ACM DL extract. The participant had to then select the article-type and run the task so that the system could retrieve the shortlisted papers. The screenshot of the task 3 from the Rec4LRW system is provided in Figure 1.

In addition to the basic metadata, the system provides the feature “View papers in the parent cluster” for the participant to see the cluster from which the paper has been shortlisted. The evaluation screen was provided to the user at the bottom of the screen (not shown in Figure 1). The participants had to answer seven mandatory survey questions and one optional subjective feedback question as a part of the evaluation. The seven survey questions and the corresponding measures are provided in Table 1. A five-point Likert scale was provided for measuring participant agreement for each question. The measures were selected based on the key aspects of the task. The measures *Relevance*, *Usefulness*, *Importance*, *Certainty*, *Good\_List* and *Improvement\_Needed* were meant to ascertain the quality of the recommendations. The final measure *Shortlisting\_Feature* was used to identify whether participants would be interested to use this task in current academic search systems and digital libraries.

**Table 1.** Evaluation measures and corresponding questions

Measure	Question
Relevance	The shortlisted papers are relevant to my article-type preference
Usefulness	The shortlisted papers are useful for inclusion in my manuscript



Importance	The shortlisted papers comprises of important papers from my reading list
Certainty	The shortlisted list comprises of papers which I would definitely cite in my manuscript
Good_List	This is a good recommendation list, at an overall level
Improvement_Needed	There is a need to further improve this shortlisted papers list
Shortlisting_Feature	I would like to see the feature of shortlisting papers from reading list based on article-type preference, in academic search systems and databases

The response values ‘Agree’ and ‘Strongly Agree’ were the two values considered for the calculation of agreement percentages for the evaluation measures. Descriptive statistics were used to measure central tendency. Independent samples t-test was used to check the presence of statistically significant difference in the mean values of the students and staff group, for the testing the hypothesis. Statistical significance was set at  $p < .05$ . Statistical analyses were done using SPSS 21.0 and R. Participants’ subjective feedback responses were coded by a single coder using an inductive approach [1], with the aim of identifying the central themes (concepts) in the text.

The study was conducted between November 2015 and January 2016. Out of the eligible 230 participants, 116 participants signed the consent form and completed the whole study inclusive of the three tasks in the system. 57 participants were Ph.D./Masters students while 59 were research staff, academic staff and librarians. The average research experience for Ph.D. students was 2 years while for staff, it was 5.6 years. 51% of participants were from the computer science, electrical and electronics disciplines, 35% from information and communication studies discipline while 14% from other disciplines.

## 6 Results and Discussion

### 6.1 Agreement Percentages (AP)

The agreement percentages (AP) for the seven measures by the participant groups are shown in Figure 2. In the current study, an agreement percentage above 75% is considered as an indication of higher agreement from the participants. As expected, the AP of students was consistently higher than the staff with the biggest difference found for the measures *Usefulness* (82.00% for students, 64.15% for staff) and *Good\_List* (76.00% for students, 62.26% for staff). It has been reported in earlier studies that graduate students generally look for assistance in most stages of research [9]. Consequently, students would prefer technological interventions such as the current system due to the simplicity in interaction. Hence, the evaluation of students was evidently better than staff. The quality measures *Importance* (85.96% for students, 77.97% for staff) and *Shortlisting\_Feature* (84.21% for students, 74.58% for staff) had the highest APs. This observation validates the usefulness of the technique in identifying popular/seminal papers from the reading list. Due to favorable APs for the most measures, the lowest agreement values were observed for the measure *Improve-*

*ment\_Needed* (57.89% for students, 57.63% for staff). The results for the measure *Certainty* (70% for students, 62.26% for staff) indicate some level of reluctance among the participants in being confident of citing the papers. Citation of a particular paper is subject to the particular citation context in the manuscript, therefore not all participants would be able to prejudge their citation behavior. In summary, participants seem to acknowledge the usefulness of the task in identifying important papers from the reading list. However, there is an understandable lack of inclination in citing these papers. This issue is to be addressed in future studies.

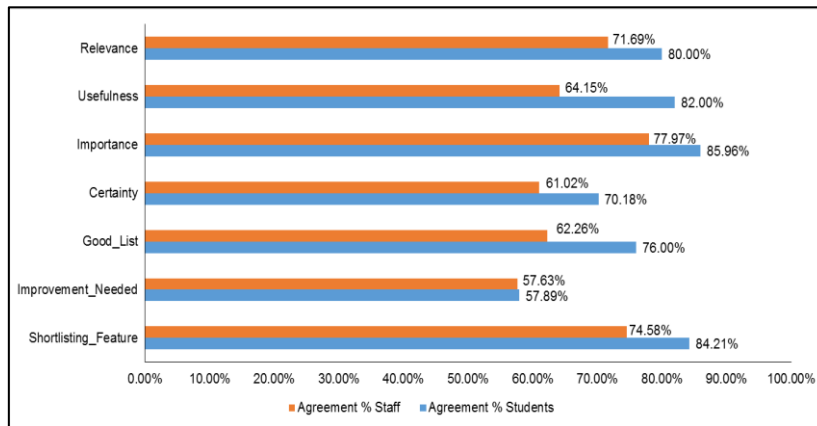


Fig. 2. Agreement percentage results by participant group

## 6.2 Qualitative Data Analysis

In Table 2, the top five categories of the preferred aspects and critical aspects are listed.

**Preferred Aspects.** Out of the total 116 participants, 68 participants chose to give feedback about the features that they found to be useful. 24% of the participants felt that the feature of the shortlisting papers based on article-type preference was quite preferable and would help them in completing their tasks in a faster and efficient manner. They also felt that the quality of the shortlisting papers was satisfactory. 15% of the participants felt that the information cue labels (popular, recent, high reach and literature survey) were helpful for them in relevance judgement of the shortlisted papers. This particular observation of the participants was echoed for the first two tasks of the Rec4LRW system, thereby validating the usefulness of information cue labels in academic search systems and digital libraries. Around 11% of the participants felt the option of viewing papers in the parent cluster of the particular shortlisted papers was useful in two ways. Firstly, it helped in understanding the different clusters formed with the references and citations of the papers in the reading list. Secondly, the clusters served as an avenue for finding some useful and relevant papers in serendipitous manner as some papers could have been missed by the researcher dur-

ing the literature review process. The other features that the participants commended were the metadata provided along with the shortlisted papers (citations count, article summary) and the paper management collection features across the three tasks.

**Table 2.** Top five categories for preferred and critical aspects

Rank	Preferred aspects categories	Critical aspects categories
1	Shortlisting Feature & Rec. Quality (24%)	Rote Selection of Papers (16%)
2	Information Cue Labels (15%)	Limited Dataset Issue (5%)
3	View Papers in Clusters (11%)	Quality can be Improved (5%)
4	Rich Metadata (7%)	Not Sure of the Usefulness of the Task (4%)
5	Ranking of Papers (3%)	UI can be Improved (3%)

**Critical Aspects.** Out of the 116 participants, 41 participants gave critical comments about the task and features of the system catering to the task. Around 16% of the participants felt that the study procedure of adding 30 papers to the reading list as a precursor for running the task was uninteresting. The reasons cited were the irrelevance of some of the papers to the participants as these papers had to be added just for the sake of executing the task while some participants felt that the 30 papers count was too much while some could not comprehend why these many papers had to be added. Around 5% of the participants felt that the study experience was hindered by the dataset not catering to recent papers (circa 2012-2015) and the dataset being restricted to computer science related topics.

Another 5% of the participants felt that they shortlisting algorithm/technique could be improved to provide a better list of papers. A section of these participants needed more recent papers in the final list while others wanted papers specifically from high impact publications. Around 4% of the participants could not find the usefulness of the task in their work. They felt that the task was not beneficial. The other minor critical comments given by the participants were the ranking of the list could be improved, the task execution speed could be improved and more UI control features could be provided, such as sorting options and free-text search box.

## 7 Conclusion and Future Work

For literature review and manuscript preparatory related tasks, the gap between novices and experts in terms of task knowledge and execution skills is well-known [15]. A majority of the previous studies have brought forth assistive systems that focus heavily on LR tasks, while only a few studies have concentrated on approaches for helping researchers during manuscript preparation. With the Rec4LRW system, we have attempted to address the aforementioned gap with a novel task for shortlisting articles from researcher’s reading list, for inclusion in manuscript. The shortlisting task makes use of a popular community detection algorithm [10] for identifying communities of papers generated from the citations network of the papers from the

reading list. Additionally, we have also tried to vary shortlisted papers count by taking the article-type choice into consideration.

In order to evaluate the system, a user evaluation study was conducted with 116 participants who had the experience of writing research papers. The participants were instructed to run each task followed by evaluation questionnaire. Participants were requested to answer survey questions and provide subjective feedback on the features of the tasks. As hypothesized before the start of the study, students evaluated the task favorably for all measures. There was high level of agreement among all participants on the availability of important papers among the shortlisted papers. This finding validates the aim of the task in identifying the papers that manuscript reviewers would expect to be cited. In the qualitative feedback provided by the participants, majority of the participants preferred the idea of shortlisting papers and also thought the output of the task was of good quality. Secondly, they liked the information cue labels provided along with certain papers, for indicating the special nature of the paper. As a part of critical feedback, participants felt that the study procedure was a bit longwinded as they had to select 30 papers without reading them, just for running the task.

As a part of future work, the scope for this task will be expanded to bring in more variations for the different article-type choices. For instance, research would be conducted:- (i) to ascertain the quantity of recent papers to be shortlisted for different article-type choices, (ii) include new papers in the output so that the user is alerted about some key paper(s) which could have been missed during literature review, (iii) provide more user control in the system so that the user can select papers as mandatory to be shortlisted and (iv) Integrate this task with the citation context recommendation task [11, 14] so that the user can be fully aided during the whole process of manuscript writing.

**Acknowledgements.** This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its International Research Centres in Singapore Funding Initiative and administered by the Interactive Digital Media Programme Office.

## References

- [1] A General Inductive Approach for Analyzing Qualitative Evaluation Data: [https://flexiblelearning.auckland.ac.nz/poplh1th701/8/files/general\\_inductive\\_approach.pdf](https://flexiblelearning.auckland.ac.nz/poplh1th701/8/files/general_inductive_approach.pdf). Accessed: 2016-04-07.
- [2] Atanassova, I. and Bertin, M. 2014. Faceted Semantic Search for Scientific Papers. *PLoS Biology*. 2, (2014).
- [3] Beel, J. et al. 2013. Introducing Docear's Research Paper Recommender System. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL)* (2013).
- [4] Beel, J. et al. 2013. Research Paper Recommender System Evaluation: A Quantitative Literature Survey. *Proceedings of the Workshop on Reproducibility and Replication in Recommender Systems Evaluation (RepSys) at the ACM Recommender System conference* (2013).
- [5] Björk, B.-C. and Hedlund, T. 2003. Scientific Publication Life-Cycle Model (SPLC). *ELPUB*. (2003).

- [6] Champiri, Z.D. et al. 2015. A systematic review of scholar context-aware recommender systems. *Expert Systems with Applications*. 42, 3 (2015), 1743–1758.
- [7] Chou, J.-K. and Yang, C.-K. 2011. PaperVis: Literature Review Made Easy. *Computer Graphics Forum*. 30, 3 (Jun. 2011), 721–730.
- [8] Dehghani, Z. et al. 2011. A multi-layer contextual model for recommender systems in digital libraries. *Aslib Proceedings*. 63, 6 (2011), 555–569.
- [9] Du, J.T. and Evans, N. 2011. Academic Users' Information Searching on Research Topics: Characteristics of Research Tasks and Search Strategies. *The Journal of Academic Librarianship*. 37, 4 (Jul. 2011), 299–306.
- [10] Girvan, M. and Newman, M.E.J. 2002. Community structure in social and biological networks. *Proceedings of the national academy of sciences*. National Acad Sciences.
- [11] He, Q. et al. 2011. Citation recommendation without author supervision. *Proceedings of the fourth ACM international conference on Web search and data mining - WSDM '11* (New York, New York, USA, 2011), 755.
- [12] He, Q. et al. 2010. Context-aware citation recommendation. *Proceedings of the 19th international conference on World wide web - WWW '10*. (2010), 421.
- [13] Huang, W. et al. 2012. Recommending Citations : Translating Papers into References. *Proceedings of the 21st ACM international conference on Information and knowledge management* (2012), 1910–1914.
- [14] Huang, W. et al. 2014. RefSeer : A Citation Recommendation System. *Digital Libraries (JCDL), 2014 IEEE/ACM Joint Conference on* (2014), 371–374.
- [15] Karlsson, L. et al. 2012. From Novice to Expert: Information Seeking Processes of University Students and Researchers. *Procedia - Social and Behavioral Sciences*. 45, (Jan. 2012), 577–587.
- [16] Levy, Y. and Ellis, T.J. 2006. A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research. *Informing Science Journal*. 9, (2006).
- [17] Lu, Y. et al. 2011. Recommending citations with translation model. *Proceedings of the 20th ACM international conference on Information and knowledge management* (New York, New York, USA, 2011), 2017–2020.
- [18] Mcnee, S.M. 2006. *Meeting User Information Needs in Recommender Systems*. Proquest.
- [19] Sesagiri Raamkumar, A. et al. 2015. Rec4LRW – Scientific Paper Recommender System for Literature Review and Writing. *Proceedings of the 6th International Conference on Applications of Digital Information and Web Technologies* (2015), 106–120.
- [20] van Setten, M. et al. 2006. Goal-based structuring in recommender systems. *Interacting with Computers*. 18, 3 (May 2006), 432–456.
- [21] Shaoping, Z. 2010. *ActiveCite: An Interactive System for Automatic Citation Suggestion*.
- [22] Tarjan, R. 1972. Depth-first search and linear graph algorithms. *SIAM journal on computing*. 1, 2 (1972), 146–160.
- [23] Winoto, P. et al. 2012. Contexts in a Paper Recommendation System with Collaborative Filtering. *the International Review of Research in Open and Distance Learning*. 13, 5 (2012), 56–75.
- [24] Wu, F. and Huberman, B.A. 2004. Finding communities in linear time: a physics approach. *The European Physical Journal B-Condensed Matter and Complex Systems*. 38, 2 (2004), 331–338.
- [25] Xia, F. et al. 2013. Socially-aware venue recommendation for conference participants. *Ubiquitous Intelligence and Computing, 2013 IEEE 10th International Conference on and 10th International Conference on Autonomic and Trusted Computing (UIC/ATC)* (2013), 134–141.