

Review of Systematic Literature Review Tools

Ahmed Al-Zubidy, Jeffrey C. Carver
 Department of Computer Science
 The University of Alabama
 Tuscaloosa, AL 35487-0290
 aalzubidy@crimson.ua.edu, carver@cs.ua.edu

Keywords-Systematic Literature Review, Tools

I. INTRODUCTION

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 have been reported in the literature. For each tool, the report describes the features and limitations. To compile this data, we first searched the software engineering literature databases for papers that reported about an SLR tool or the use of a tool for conducting a SLR. After filtering the results and going through each paper we prepared a list of features and limitations for each tool and classified each tool based on the SLR stage(s) it supports. After compiling our initial information, we sent the tool reports to the tool authors for validation. We incorporated any changes returned. We used these reports as part of the SLR Community Workshop held prior to the EASE '14 conference (<http://carver.cs.ua.edu/SLR>).

The following subsections provide details about each tool identified in the literature.

II. DBPEDIA

A. Developed By

- The executive team of the DBpedia Association consists of:
 - Sebastian Hellmann
 - Michael Martin
 - Dimitris Kontokostas
 - They can be contacted at dbpedia@infai.org.
- However the paper's publishers contact information (all @polito.it):
 - Federico Tomassetti (federico.tomassetti)
 - Giuseppe Rizzo (giuseppe.rizzo)
 - Antonio Vetro (antonio.vetro)
 - Luca Ardito (luca.ardito)
 - Marco Torchiano (marco.torchiano)
 - Maurizio Morisio (maurizio.morisio)

B. Tool Features

It is designed to help in the Selection of Primary Studies stage. It is a Resource Description Framework (RDF) repository where information stored in Wikipedia is represented as structured data. This repository works as a look-up system of resources and it could be used to enrich the documents to help in the selection studies stage.

C. Limitations

The possibility that some terms are not present in DBPedia (because it depends on the Wikipedia), hence a paper will not be enrich-able and their approach cannot be applied to it

D. Reported Stages Support

Extend and Enrich the Selection of Primary Studies.

E. Papers

- Tomassetti, F., et al. (2011). Linked data approach for selection process automation in systematic reviews. Evaluation & Assessment in Software Engineering (EASE 2011), 15th Annual Conference on.
- Marshall, C. and P. Brereton (2013). Tools to Support Systematic Literature Reviews in Software Engineering: A Mapping Study. Empirical Software Engineering and Measurement, 2013 ACM / IEEE International Symposium on.

III. PROJECTION EXPLORER PEX

A. Developed By

- Dr. Rosane Minghim - rminghim@icmc.usp.br
- Dr. Maria Cristina F. Oliveira - cristina@icmc.usp.br
- Dr. Luis Gustavo Nonato - gnonato@icmc.usp.br
- Dr. Fernando Vieira Paulovich - paulovic@icmc.usp.br
- More people are available on <http://infoserver.lcad.icmc.usp.br/infovis2/People>

B. Tool Features

This tool and its extension PEx-Graph support two main stages in the SLR process, the Studies Selection stage and the Reporting or Documentation stage. The tool allows for a VTM exploration of a collection of documents on a document map, in a way that documents that are similar tend to be close together, while dissimilar documents are placed apart from each other. The VTM tool can cluster related documents and label them with terms considered more significant to represent each cluster. This feature was central for the exploration strategy devised. PEx provides functionalities to conduct preprocessing, similarity calculation and projection, creating therefore the document map as well as it has a module that processes the abstract of the primary studies, eliminates stop words, calculates the terms frequency and, based on this

result, displays clusters of studies to facilitate their analysis. All those features would support for the two stages of the SLR that were mentioned above.

C. Limitations

- Using PEx demanded a previous step in order to prepare the collected data to an acceptable format. To be explored, the document files were converted from PDF format to raw ASCII text format. The articles spreadsheet was also processed for exploring using text from parts of the articles (title, abstract, references etc). Resulting files were processed in PEx
- The choice of the VTM tool, its configuration and the required data preprocessing was not trivial requiring the intervention of the VTM specialist.
- The use of SM-VTM requires some experience and knowledge in the use of text mining and visualization tools. It is mainly required knowledge that users understand the output of projection techniques and clustering algorithms and learn how to handle word exclusion for topic determination

D. Reported Stages Support

The VTM tool was used to support articles selection.

- Categorization and classification stages in the systematic mapping using Visual Text Mining
- Studies Selection Stage
- Documentation/Reporting Stage

E. Papers

- Felizardo, K. R., et al. (2010). An approach based on visual text mining to support categorization and classification in the systematic mapping. Proc. of EASE.
- Felizardo, K. R., et al. (2011). Analysing the Use of Graphs to Represent the Results of Systematic Reviews in Software Engineering. Software Engineering (SBES), 2011 25th Brazilian Symposium on.
- Malheiros, V., et al. (2007). A Visual Text Mining approach for Systematic Reviews. Empirical Software Engineering and Measurement, 2007. ESEM 2007. First International Symposium on.
- Fabbri, S., et al. (2012). Managing Literature reviews information through visualization. 14th International Conference on Enterprise Information Systems, ICEIS 2012, June 28, 2012 - July 1, 2012, Wroclaw, Poland, International Conference on Enterprise Information.
- Marshall, C. and P. Brereton (2013). Tools to Support Systematic Literature Reviews in Software Engineering: A Mapping Study. Empirical Software Engineering and Measurement, 2013 ACM / IEEE International Symposium on.

IV. REVIS

A. Developed By

- Katia Romero Felizardo (karfelizardo@gmail.com / katiarf@icmc.usp.br)

- Rafael Messias Martins (rmartins@icmc.usp.br)
- José Carlos Maldonado (jcmaldon@icmc.usp.br)
- <http://ccsl.icmc.usp.br/pt-br/projects/revis>

B. Tool Features

ReVis is a text mining tool that supports the study selection and study selection review stages in the SLR process using content-based analysis of documents (content map) and meta-data analysis, via representations such as edge bundles and citation map. The content map enables users to investigate content similarity relationships among the pairs of data points. The other two visualizations use citations among studies and shared references between them to indicate citation relationships. There are four strategies to explore the content map: (i) clusters and topics it applies clustering algorithms in order to create clusters and their respective topics; (ii) expression occurrence it shows the frequency of occurrence of specific user-defined expressions in the primary studies; (iii) KNN edges connection (neighborhood relationship) it allows users to explore neighborhood relationships in the map, with neighboring connections (called KNN connections) shown as edges between the points; (iv) coordination links are created between the same documents in different visual representations. All those features and more of text mining features would help identifying the studies easier and would help making a decision to include or exclude a study faster. Revis also can be used to support the updating of SLRs and the categorization and classification stage in the mapping study process.

C. Limitations

- The use of Revis to support the selection and selection review activities generally requires extra time to prepare and provide the information (title, abstract, keywords and references) about the primary studies for the tool
- The time taken depends on different variables, such as: (i) the number of primary studies to be analyzed; (ii) the literature search used, for example, if the Web search engines provided by digital libraries are or not automated search, the number of manual search conducted, among others;
- If the studies are in any other format (e.g. PDF), it is necessary to convert them prior to the analysis.

D. Reported Stages Support

Study Selection and Study Selection Review Stages (SLR).

E. Papers

- Felizardo, K.R., et al. (2010) An Approach Based on Visual Text Mining to Support Categorization and Classification in the Systematic Mapping. International Conference on Evaluation and Assessment in Software Engineering (EASE).
- Felizardo, K. R., et al. (2011). Using Visual Text Mining to Support the Study Selection Activity in Systematic Literature Reviews. International Symposium on Empirical Software Engineering and Measurement (ESEM).

- Felizardo, K. R., et al. (2012). A Visual Analysis Approach to Validate the Selection Review of Primary Studies in Systematic Reviews. *Information and Software Technology* 54(10): 1079-1091.
- Felizardo, K.R.; et al. (2013) The Use of Visual Text Mining to Support the Study Selection Activity in Systematic Literature Reviews: A Replication Study. *International Workshop on Replication in Empirical Software Engineering Research (RESER)*.
- Felizardo, K.R., et al. (2013) A Visual Approach to Validate the Selection Review of Primary Studies in Systematic Reviews: A Replication Study. *International Conference on Software Engineering & Knowledge Engineering (SEKE)*.
- Felizardo, K.R.; et al. (2013) A Visual Analysis Approach to Update Systematic Reviews. *International Conference on Evaluation and Assessment in Software Engineering (EASE)*.
- Marshall, C. and P. Brereton (2013). Tools to Support Systematic Literature Reviews in Software Engineering: A Mapping Study. *International Symposium on Empirical Software Engineering and Measurement (ESEM)*.

V. SLURP

A. Developed By

- David Bowes - d.h.bowes@herts.ac.uk
- Tracy Hall - tracy.hall@brunel.ac.uk
- Sarah Beecham - sarah.beecham@lero.ie

B. Tool Features

- Selection of Primary Studies - SLuRP can apply pre-defined search terms to online databases (not all the databases because it is rejected by the database to use automation tools) and semi-automatically extract papers from databases and save them as well as extract and save their PDFs. SLuRP prompts users to assign two or more reviewers to each paper, then it records the evaluation of each reviewer alongside the inclusion/exclusion criteria. If there is a need to settle disagreements about studies between reviewers, it will be flagged by SLuRP. "The frequency of disagreements for including or excluding papers is automatically generated by SLuRP. This allows inter-rater reliability scores to be easily produced".
- Assess Study Quality - SLuRP allows the reviewers to define a group of quality criteria. It allows reviewers to store all data extracted from included papers which have approved the quality check. It can store data about the study and the context of the study. It ensures the validity to the data extracted from included papers.
- Synthesize Data - SLuRP aggregates quantitative conclusions and displays these in tables, or graphical form. It allows the results of the data analysis to be presented in two forms, either tabular or graphical. These graphical display can emphasize trends in results and allow the researchers to consider bias in results. SLuRP supports the researchers to synthesize qualitative data, based on the research questions, by storing or keeping qualitative

information which can then be analyzed in tabular and graphical format

C. Limitations

No limitations were reported because this paper was by the developers of the tool.

D. Reported Stages Support

The entire SLR process.

E. Papers

- Bowes, D., et al. (2012). SLuRP - A tool to help large complex systematic literature reviews deliver valid and rigorous results. *2nd International Workshop on Evidential Assessment of Software Technologies, EAST 2012, September 22, 2012 - September 22, 2012, Lund, Sweden, Association for Computing Machinery*.
- Marshall, C. and P. Brereton (2013). Tools to Support Systematic Literature Reviews in Software Engineering: A Mapping Study. *Empirical Software Engineering and Measurement, 2013 ACM / IEEE International Symposium on*.

VI. START

A. Developed By

Federal University of São Carlos. http://lapes.dc.ufscar.br/tools/start_tool

- Sandra Fabbri - sfabbri@dc.ufscar.br
- Fabiano Ferrari - fabiano@dc.ufscar.br

B. Tool Features

This tool could be used for all stages of the SLR process:

- Planning -The tool provide an interface to enter keywords that will be used to compose the search strings. When studies are uploaded into StArt, it uses the keywords to score the studies according to the number of occurrences of these words in their title, abstract and keywords. It also provide an interface to enter the inclusion and exclusion criteria which will be used to either accept or reject a study. Finally, in this stage it provides another interface to enter information about the extraction form which must be completed during the Extraction stage.
- Execution - The researcher will upload the BibTex file to the StArt tool. Selection: In this step the primary studies uploaded into StArt must be accepted or rejected according to the inclusion and exclusion criteria defined in the first stage. The decision should be made after reading the title, abstract and keywords of the study. In the Extraction step, all studies that have been accepted during the Selection step should be read in full and be analyzed again. The researcher must complete the Data Extraction Form defined during the Planning stage.
- Summarization - This tool facilitates the access to the information extracted during the Extraction step and provides a text editor to help in a first version of the

summarization document. When this stage is reached, some data on the whole SLR is available, StArt provides some reports that also facilitate the conduction of a SLR. The Start tool provides a visualization of all the studies retrieved as well as their references. Hence, it is possible to identify for example, if a frequently cited reference was or was not retrieved by the search string. Finally it helps the generation of charts that support a quantitative SR characterization

C. Limitations

- The researcher must do the search manually through the search engines registered in the Protocol every time a search is necessary. The search result must be exported from the search engine as a BibTex file which must be imported into StArt
- The strings should be applied in each search engine, for example, IEEE, Scopus, ACM, Springer and Web of Science. This action is not supported by the tool and the search results must be imported into StArt
- (Data Extraction Phase) The researcher must read the full version of each study Accepted study, elaborate a summary and fill in the Information Extraction Form of each study (Figure 4-B). Aiming to facilitate this step, it is possible to link the studies full text file (e.g. PDF files) with their record in the tool.

D. Reported Stages Support

All the stages

E. Papers

- Fabbri, S., et al. (2012). Managing Literature reviews information through visualization. 14th International Conference on Enterprise Information Systems, ICEIS 2012, June 28, 2012 - July 1, 2012, Wroclaw, Poland, International Conference on Enterprise Information.
- Hernandez, E., et al. (2012). "Using GQM and TAM to evaluate StArt-a tool that supports Systematic Review." CLEI Electronic Journal 15(1): 3-3.
- Hernandez, E. M., et al. (2013). Experimental studies in software inspection process: A systematic mapping. 15th International Conference on Enterprise Information Systems, ICEIS 2013, July 4, 2013 - July 7, 2013, Angers, France, SciTePress
- Marshall, C. and P. Brereton (2013). Tools to Support Systematic Literature Reviews in Software Engineering: A Mapping Study. Empirical Software Engineering and Measurement, 2013 ACM / IEEE International Symposium on.