Challenges and Pitfalls on Surveying Evidence in the Software Engineering Technical Literature: an Exploratory Study with Novices

Talita Vieira Ribeiro
PESC/COPPE
Federal University of Rio de Janeiro
Rio de Janeiro, RJ, Brazil
tvribeiro@cos.ufrj.br

Jobson Massollar
PESC/COPPE
Federal University of Rio de Janeiro
Rio de Janeiro, RJ, Brazil
jobson@cos.ufrj.br

Guilherme Horta Travassos
PESC/COPPE
Federal University of Rio de Janeiro
Rio de Janeiro, RJ, Brazil
ght@cos.ufrj.br

1 INTRODUCTION

The evidence-based software engineering approach advocates the use of scientific evidence by software engineers to support the adoption of software technologies in industrial software development and maintenance projects. Aside from the unavailability of scientific knowledge in industrial settings and the time required to acquire evidence in the software engineering (SE) field, additional challenges prevent practitioners, mainly those that are not experienced in research, from collecting knowledge from scientific sources to support the decision-making throughout their software projects.

2 THE EXPLORATORY STUDY

In [1], we identify pitfalls and characterize challenges on surveying evidence in the SE technical literature. Our observations are supported by observing similarities and differences in the planning and generated results of systematic literature review (SLR) research protocols. These secondary studies regard the same research question (quality attributes of use cases) and similar teams of novice researchers (with some experience in development but without research experience, as it usually happens with practitioners in the software industry) undertook them.

We qualitatively compared (using Jaccard and Kappa coefficients) and evaluated (using an adapted version of DARE) the results of seven SLR research protocols performed in two distinct years: 2010 and 2012. The search strings; concepts abstracted from the search terms; inclusion and exclusion criteria; returned, included and excluded papers; and quality attributes were used as units of analysis for making the comparison among the teams' plans and reports.

3 RELEVANT OUTCOMES

We observed unexpected differences in the seven SLRs' planning and execution. Even when the research protocols reached some planning agreement; the outcomes were different. Among the main observed pitfalls, the ones related to unnecessary and inappropriate search terms, lack of transparency

of decisions regarding the strategy to select papers, and misinterpretation of the research questions influenced the most the differences presented in Figure 1. Six challenges on surveying evidence in SE were characterized from such pitfalls: researchers' inexperience in the topic, researchers' inexperience in the method, lack of papers clearness and completeness, lack of a common problem domain terminology, lack of research verification procedures, and lack of commitment to the SLR.

	# search terms	# search concepts	# returned papers	# included papers	# excluded papers	# quality attributes
	11	7	661	20	641	22
	67	13	521	27	494	29
	39	9	157	19	138	19
	215	21	399	14	385	27
	40	10	351	6	345	17
	62	7	643	7	636	27
	81	8	157	16	141	27

Figure 1: Summary of the teams' quantitative results.

4 CONCLUSIONS

According to our findings, it is not possible to rely on results of SLRs performed by novices. As an example, from the 84 quality attributes reported by the seven teams, only five attributes were reported by all teams. Moreover, non-explicit information might entail differences in the outcomes even in the face of similarities in SLR plans, hampering the repeatability and confidence of the SLR process and results. While problems related to a standard SE terminology and the completeness of scientific reports are not solved, systematic planning, transparency of decisions and verification procedures are vital factors to guarantee the reliability of SLRs in our field.

REFERENCES

[1] Talita V. Ribeiro, Jobson Massollar and Guilherme H. Travassos. 2017. Challenges and pitfalls on surveying evidence in the software engineering technical literature: an exploratory study with novices. *Empirical Software Engineering*, 1-70. DOI: https://doi.org/10.1007/s10664-017-9556-7 Available at: http://rdcu.be/xNku