# An Evaluation of Knowledge Translation in Software Engineering


Deepika Badampudi
*Department of software engineering*
*Blekinge Institute of Technology*
Karlskrona, Sweden
deepika.badampudi@bth.se

Claes Wohlin
*Department of software engineering*
*Blekinge Institute of Technology*
Karlskrona, Sweden
claes.wohlin@bth.se

Tony Gorschek
*Department of software engineering*
*Blekinge Institute of Technology*
Karlskrona, Sweden
tony.gorschek@bth.se



***Abstract*—Knowledge translation is defined, in health sciences, as “the exchange, synthesis and ethically sound application of research results in practice”. The objective of this paper is to implement and conduct a feasibility evaluation of a knowledge translation framework in software engineering. We evaluated the outcome of the knowledge translation framework in an industrial setting, along with the effectiveness of the interventions undertaken as part of knowledge translation in a multi-case study. The results of the evaluation suggest that the practitioners perceive the knowledge translation framework to be valuable and useful. In conclusion, this paper contributes towards the reporting of a systematic implementation of knowledge translation and evaluating its use in software engineering.**



***Index Terms*—Knowledge translation, technology transfer, systematic literature reviews**


## I. Introduction

The research, presented here, builds on the vision that research results should be translated into actionable real-world guidance [1]. However, software practitioners often rely on their experiences when making decisions [2]. To facilitate informed decisions, there is a need to translate knowledge to software practice [3]. Knowledge translation (KT) is done in an ad-hoc manner in software engineering [3]. Thus, this paper contributes with documenting a feasibility implementation and evaluation of a KT framework for software engineering. The current work is intended as a stepping stone towards further more complete evaluations of the KT framework, which has been recognized as a need [4].

An overview of the KT framework is depicted in Figure 1. The KT framework as presented in [4], consists of Bayesian synthesis and the Plan-Do-Check-Act (PDCA) improvement paradigm as shown in Figure 1. The Bayesian synthesis [5] includes prior probability (prior opinions/experiences of practitioners), likelihood (what is known in literature, i.e. research results) and posterior probability (revised prior in the light of the likelihood). Bayesian synthesis (Steps 1-4 in Figure 1) helps in contextualizing the research results in practice and PDCA improvement paradigm (Steps 5-9 in Figure 1) helps in implementing the knowledge in practice. The KT framework consists of nine steps. In summary, the KT steps are as follows:



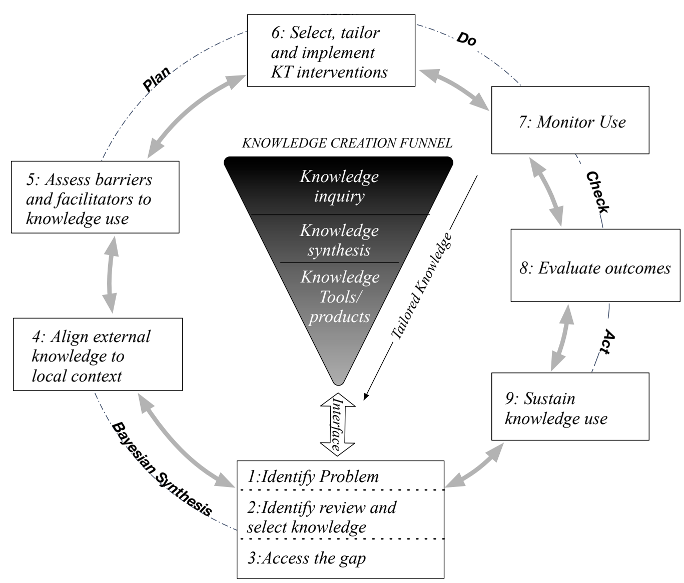


Fig. 1. Knowledge translation framework (Adapted from Graham et al. [6]).

1) Identify problem: Elicit prior probability - In this step, the relevant practitioners associated with the improvement initiative/need are selected and their prior beliefs based on their experiences are elicited.
2) Identify, review and select knowledge - The knowledge to be translated is identified, reviewed and selected from the literature. The identified knowledge should be classified based on the different levels of evidence [1].
3) Assess gaps - The gap between current and desired knowledge is assessed in this step.
4) Align external knowledge to local context - The knowledge from research studies (identified and classified in Step 2) is presented to the practitioners as likelihood and their prior beliefs (elicited in Step 1) are revised in the light of the likelihood.
5) Assess barriers and facilitators to knowledge use - The barriers that could potentially prevent the knowledge use and facilitators that could promote knowledge use in practice are assessed.
6) Select, tailor and implement the KT intervention - Based on the barriers and facilitators (identified in Step 5), the

interventions (for example, workshops or trainings) for knowledge use in a particular context are decided and implemented.

7) Monitor use - In this step, the use of translated knowledge in a particular context is monitored.
8) Evaluate outcomes - The outcome of knowledge use in a context as well as, the KT intervention (implemented in Step 6) is evaluated.
9) Sustain use - Sustainability-oriented action plans to continue the knowledge use after the translation is identified and selected in this step.

The objective of this paper is to present a first feasibility evaluation of the KT framework introduced in [4]. The details of each step and its implementation are provided in Section III. We conducted a multi-case study to evaluate the first eight steps of the KT framework as shown to the left in Figure 2. The ninth step of the KT framework is more applicable when the use of knowledge is instrumental, for example, introducing a new process. In the evaluation case, the use of knowledge was on a conceptual level, i.e. use of knowledge to support the software practitioners understanding. Thus, the ninth KT step is not considered in this evaluation.

The research methods used in the evaluation as shown to the right in Figure 2 and the details of the evaluation case are presented in Section II. The details of the KT framework implementation and evaluation are provided in Section III. The evaluation results are discussed in Section IV and conclusions are presented in Section V.

## II. Research Method

In this section, we discuss the research steps taken to implement and evaluate the KT framework. The research steps undertaken and their timeline are presented to the right in Figure 2.

Case studies investigate a contemporary phenomenon within its real-life context [7]. Since, the aim was to translate knowledge from literature to practice, we used a case study to understand the state of practice and align knowledge from literature to the local context.

Two cases were selected in the form of two subsystems from a case organization that largely focuses on providing telecommunications services. They were interested in improving their GUI testing through automation. Therefore, the need to translate knowledge from existing research studies was identified.

The duration of the evaluation study was one year from starting the work in improving test automation to evaluation of the KT framework as a supportive method to facilitate integration of research results into the improvement. The improvements at the case organization were done by master students who worked approximately for six months in the case organization. The focus in this paper is on the evaluation of the KT framework, which was provided to the master students for use as part of their master thesis work. The authors of this paper implemented the KT interventions and conducted its evaluation, which was not part of the master thesis work.

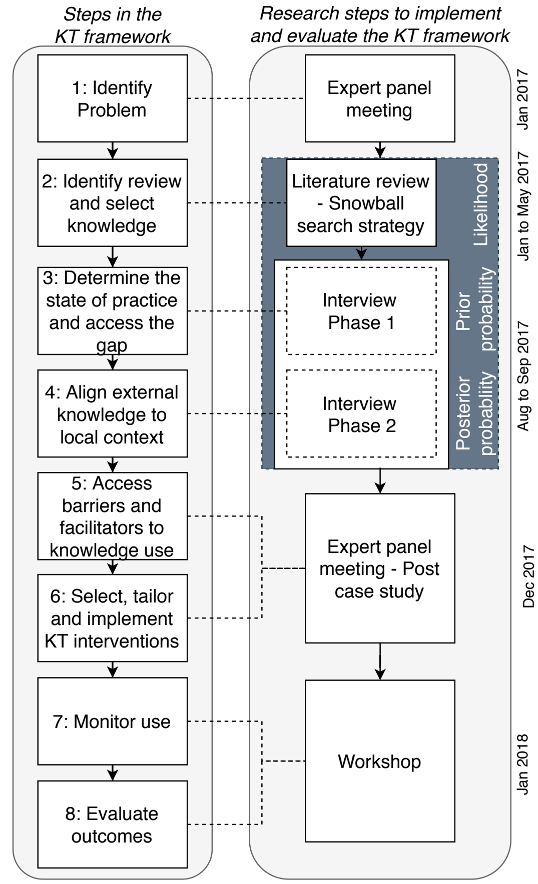


Fig. 2. Research steps undertaken to implement and evaluate the KT framework.

The description of the research steps executed is as follows: **Expert panel meeting** - As shown in Figure 2, the area for the master thesis was identified in an expert panel meeting wherein the first two authors, a master student and a local champion from the case organization collectively discussed and decided the scope of the knowledge translation, i.e. GUI test automation. The objective was to understand what are the *criteria* for when to begin automation? What are the *testability* requirements for performing GUI test automation? And what are the factors that lead to *waste* in GUI test automation? Waste is described as the additional effort spent in maintaining the GUI test scripts which could be avoided when tests are written for reuse.

**Literature review** - A literature review was conducted to identify, review and select knowledge from research studies relevant to the criteria, testability and waste in GUI automation. Snowballing was used as a search strategy to find relevant papers using the guidelines in [8]. Two masters students conducted the literature review while the first author of this paper reviewed the search protocol, inclusion and exclusion criteria, data extraction scheme and the analysis approach. In total, 19 papers were selected. As the main focus of the paper is on the KT framework and its evaluation, we have not discussed the literature review protocol and the literature review results in this paper.

**Interviews** - The local champion provided a list of all employees working with the two subsystems. From the given list, we used cluster sampling for selecting the interview participants. The aim was to select at least one person randomly from each role involved in GUI testing in both subsystems. In addition, participants were selected based on their availability. The master students developed the interview protocol and conducted the interviews. The first author reviewed the interview protocol. Each interview lasted for 90 minutes. The interviews were divided in two phases as shown in Figure 2.

In **Phase 1**, generic questions related to the role of the interviewee along with the description of the GUI test procedure were asked. In addition, questions concerning criteria for when to automate, testability requirements and factors related to waste in GUI test automation were asked. Phase 1 of the interview was focused on eliciting the opinions and experience of the interviewees. Thus, it is related to the prior probability step of Bayesian synthesis.

In **Phase 2** of the interview, the knowledge from the research results were shared with the interviewees. In addition, their revised opinions in the light of knowledge from research studies was elicited (posterior probability).

**Expert panel meeting - post case studies** - We conducted a meeting with the local champion after all the interviews were conducted. The first two authors of the paper lead the meeting. We presented the results from the literature review and the opinions of interviewees in the form of prior and posterior probabilities to the local champion. The goal of the meeting was to collectively decide the next steps for knowledge implementation if needed. As the knowledge translation was done on a individual level through the interviews, conducting a workshop to discuss the results and attain consensus was decided.

**Workshop** - The first two authors conducted the workshop with the interviewees and the local champion. The agenda of the workshop was to first discuss and reflect on the interview results, monitor knowledge use and evaluated the knowledge use outcomes. To facilitate the discussions we asked the following semi-structured questions:

1) Are the interviewees' opinions aligned with each other and with the literature review results?
2) The interviewees did not mention some of the literature review results in Phase 1 of the interview, did/would you introduce these criteria, testability requirements and waste in GUI automation? Or did/would it change the way you think? Or did it affect your knowledge levels, understanding or attitude?

We also collected feedback on the entire study procedure, i.e., the KT framework implementation, by asking the following questions:

1) What would you change in the interview or workshop procedure?
2) What worked and what did not work well in the interview or workshop?
3) Was the research study valuable to you?
4) Was the research study an important utility?
5) Was information about research (potentially at other companies) important?

The results of the KT framework implementation and evaluation is discussed in Section III

## III. KT FRAMEWORK IMPLEMENTATION AND EVALUATION

Case description - The KT framework was used to translate evidence from research studies related to three areas:

- What are the *criteria (C)* for when to begin automation?
- What are the *testability (T)* requirements for performing GUI test automation?
- What is *waste (W)* in test automation?

In particular, the unit of analysis were as follows - *criteria* for when to automate, *testability* requirements and factors associated with *waste*. The outcome of the knowledge translation with respect to the criteria, testability requirements and waste in GUI automation is depicted in Figures 3, 4 and 5 respectively and elaborated in Section III-A.

**Need for knowledge implementation:** The expert panel consisting of a test expert (local champion), a professor, a PhD student and a master student identified the need for improvement in terms of alignment in the practitioners' opinions and knowledge related to GUI test automation. The test expert was the local champion, in the case organization, and he was interested to know the perceptions of the team members as well as the evidence available from research studies. The goal, as described by the local champion, was to have a common understanding regarding the criteria, testability requirements and waste.

### *A. The steps followed in the KT framework*

*1) Identify problem - Elicit prior probabilities:* **Selecting individuals:** We selected the practitioners related to the testing activity i.e., test lead, test architect, developer, tester, design lead and product owner. In this particular case, all roles involved in the project were selected to understand the different perceptions. In total, 12 interviews were conducted representing the two subsystems.

**Elicit opinions:** The state-of-practice in terms of prior probabilities are elicited through interviews from the above mentioned stakeholders. The practitioners were asked about their testing process. In particular, the following questions were asked: What criteria do you use to decide when to automate? What testability requirements do you consider? What are the factors that you associate with waste? The prior probabilities in the Bayesian synthesis is set to be equal to the frequencies. The frequencies are in the form of percentages of the interviewees mentioning any one of the areas related to the criteria, testability and waste in GUI automation. For example, if a criterion for when to automate was mentioned by five out of 12 interviewees then, the prior probability is 42%. The prior probability is indicated by dotted lines (....) in Figures 3, 4 and 5.

**Observations on the prior probability**: Not all interviewees mentioned the same criteria, testability requirements and waste in GUI automation. For example, only criterion C2 in Figure 3 was mentioned by all interviewees. In most cases only a subset of the interviewees identified the criteria. Out of the total number of possible responses (12 interviewees x 10 criteria = 120) from interviewees only 42% of the interviewees' responses accounted for the identified criteria in Figure 3. This indicates that all stakeholders are not equally informed or are not able to articulate or communicate the same criteria when being asked. This further implies that when decisions are made or processes are defined, some of the criteria might not be considered or communicated.

*2) Identify review and select knowledge (Likelihood):* A literature review was conducted to identify relevant knowledge from research studies. Overall, we found 19 related studies and we extracted the frequency and qualitative information related to the criteria, testability requirements and waste from the identified studies. Similar to the prior probability calculations, the percentage of the research studies mentioning the criteria, testability requirements and waste in GUI automation were calculated. The likelihood is indicated by dashed lines (- - - -) in Figures 3, 4 and 5. In addition, we extracted qualitative information about each criterion, testability requirements and waste. Qualitative information includes the name, description and references of the papers that mention the criterion, testability requirements and factors associated with waste.

**Observations on the Likelihood** - The likelihood calculation are based on the percentage of papers that discuss a particular finding. The papers might focus on a particular subset of findings and not all possible findings. Therefore, a 10% likelihood does not necessarily mean that only 10% of the papers found the finding. It merely indicates the research interest in a specific finding.

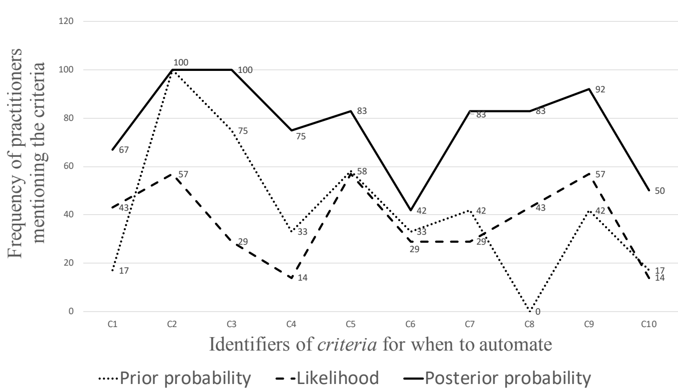


Fig. 3. Criteria for when to automate

*3) Assess the gap:* The local champion wanted to know the team members' perceptions and also about the evidence in the research studies. In addition, the desired state was that all team members shared the same knowledge.

*4) Align external knowledge to local context:* The knowledge identified, reviewed and selected in Step 2 is presented to the interviewees and their opinions (posterior probabilities) were elicited in the light of knowledge that was presented to them. The posterior probability is indicated by solid lines (—) in Figures 3, 4 and 5.

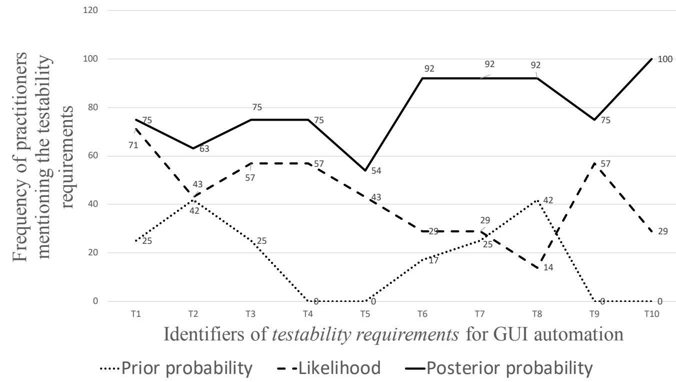


Fig. 4. Testability criteria

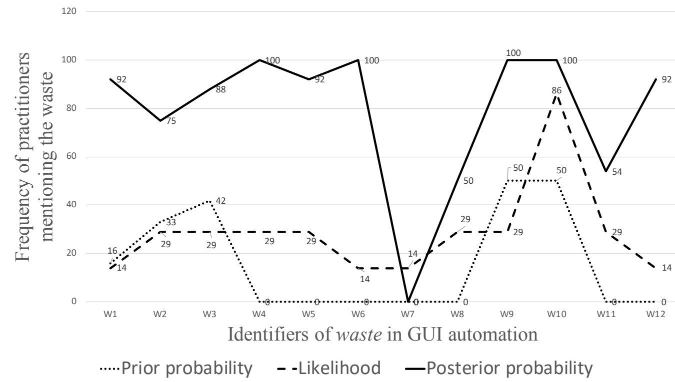


Fig. 5. Factors associated with waste

**Observations on the posterior probability**: Posterior probability is not always 100%, though the results show that the respondents became more consensual, it still shows that complete consensus is not achieved. In some cases, the likelihood (in terms of studies in the literature) is not applicable or aligned to the context hence, there is no difference between the posterior probability and the prior probability (For example, Waste W7 in Figure 5).

For the criteria, testability requirements and waste in GUI automation mentioned in Phase 1 of the interview (prior probability) and reported in the literature, the average posterior probability was 81% whereas the average prior probability was only 39%. Which indicates that after the knowledge translation the interviewees became more consensual and more informed by having the knowledge available from literature.

Even when the interviewees did not mention the criteria, testability requirements and waste in GUI automation in their prior probability, the final probability increased drastically. The average posterior probability was 73% even when the likelihood percentage was rather low, i.e. 26% on an average. If the research result is applicable to the context then the interviewees revised their opinions, i.e. posterior probability increased however, the research that is applicable to a specific context might not be heavily investigated. Therefore, it is not surprising that there was a significant increase in the posterior probability despite the likelihood being low.

*5) Assess barriers and facilitators:* The knowledge was translated through one-on-one interviews. The possible barrier anticipated with the interviews was that the interviewees might tend to agree more to avoid discussions or due to time constraints. On the other hand, the interviewees might not agree or not know about certain criteria for example that are related to other interviewees with different roles or tasks. In other words, the interviewees' opinions and knowledge might not be aligned.

*6) Select, tailor and implement KT interventions:* Due to the identified barriers and the need to have a joint discussion, we decided to conduct a workshop with the interviewees. We decided to tailor the workshop by only using aggregated results and not asking individual questions on why they did not mention certain criteria, testability requirements and waste in GUI automation. Rather than discussing what they did not mention in the interview we discussed what other companies (e.g. results of case studies reported in the literature) and researchers think and if they agree or disagree to others. This format of the workshop encourages discussion rather than focusing on individual responses. In addition, the intention of conducting a workshop was to facilitate discussions among workshop participants that had different roles to understand the big picture of the end-to-end test automation process.

*7) Monitor use:* Since, the identified need for improvement was to align the practitioners' knowledge and opinions, the knowledge use was on the conceptual level. Therefore, in the workshop we collected information about the levels of knowledge and attitudes before and after being informed about the research results.

We asked about the criteria, testability requirements and waste in GUI automation that were not mentioned by all the interviewees. Based on the response of the workshop participants, we conclude that the practitioners' opinions before knowledge translation are driven by a number of factors: the tasks practitioners are involved in at the time of the interview, their role and the background such as education and work experience. If practitioners miss out on mentioning a criterion, testability requirement or waste in GUI automation, then it might be due to the fact that it is not a priority for the practitioner working on a particular task at the time of the interview. For example, as a tester, the task was to automate everything, so when asked about criteria for when to automate, the tester is likely to mention the criteria that are relevant to his/her task. Criterion such as feasibility assessment of test automation might be a criterion that is considered by the tester and most likely the he/she will not mention it right away.

In addition, the question regarding the intention to change based on the knowledge from research results were asked. The workshop participants were already following most of the practices mentioned in the literature. However, we found some new insights which were not considered by the practitioners prior to the KT framework implementation. However, the practitioners did not perceive them to have much impact. Although, this is a reflection on the results from research studies and not so much on the KT framework itself.

*8) Evaluate outcomes:* Several perspectives are important to evaluate. The need for improvement was on the practitioner level, therefore, we collected the practitioner's satisfaction, confidence and knowledge through the workshop. The evaluation of the KT outcome is based on the time, usefulness and value of KT as perceived by the workshop participants.

*Time:* Practitioners hardly have time to find answers in research studies. Therefore, it was appreciated that the research results were summarized to them and it was considered to be a time efficient process. In addition, scientific results were valued more than the external information practitioners access (usually technical blogs).

*Usefulness:* The KT steps helped in aligning and streamlining the opinions of the workshop participants. They appreciated the external input and the knowledge as sometimes they can get too influenced by their own tasks and environment. They also identified the most important use of KT framework in supporting (informed) decision-making. Particularly when they do not have all the information needed to evaluate all the alternatives.

*Value:* The workshop participants expressed an increase in confidence as they got a confirmation that they are following good practices and that they are not missing any crucial knowledge from research.

*Evaluation of KT intervention:* The interventions used in the implementation of the KT framework were interviews and a workshop. The combination of interviews and workshop was appreciated. The interviews allowed the practitioners to express their individual opinions without being influenced by others. At the same time, after expressing their opinions, it was good to discuss with the other practitioners and reflect more on it. In addition to the conversation between the researchers and practitioners during the workshop, the practitioners asked questions to other practitioners and discussed among themselves as well. The practitioners did more reflecting and discussed the criteria, testability requirements and waste in GUI automation that they did not mention in the interview.

## IV. Discussions of the KT evaluation

The KT framework was positively received by the practitioners. The practitioners perceive KT to be useful in making decisions, in particular, when alternatives are not explicitly known. The practitioners expressed an increase in confidence regarding their knowledge of GUI testing.

The expert panel discussion, including key stakeholders (researchers and local champion) at the beginning of the study, was a good input. It helped the researchers to understand the industrial need and the local champion was able to communicate what is important for them. Another advantage was the possibility to interview all the roles involved in testing. It helped in understanding the topic from all perspectives.

The practitioners did not elicit the same results, i.e. only 42% of the respondents mentioned the total number of criteria mentioned in the interviews related to when to automate. The likelihood was presented in terms of percentages and qualitative information was provided during the interview. The practitioners became more consensual after knowing the research results, i.e., 39% of the respondents mentioned the overall criteria, testability requirements and waste in GUI automation before knowing the research results, which increased to 81% after knowing the research results. This indicates that the practitioners had a more common understanding among themselves after the KT framework was used. Overall, based on the evaluations, we conclude that the KT framework was successfully implemented in practice. However, we identify minor challenges that could be addressed to make the KT framework more effective. Based on the challenges, we identified the adaptations to KT framework that are discussed in Section IV-A.

### *A. Updates in the KT framework based on the challenges in the evaluation*

*The difference in terminology:* The same terminology used in the research studies was used in the interviews; it created some problems in communicating the knowledge and created problems for practitioners in interpreting the results. Hence, after the information is retrieved from the research studies in Step 2, it is important to discuss the results with the local champion. The results ought to be summarized in the terminology used in the company; this will help in aligning research findings to the local context in Step 4.

*Allowing time between collecting prior and posterior opinions:* Practitioners wanted to have time gap between collecting the prior opinions and the revised opinions. After the prior opinions were elicitation, the practitioners wanted to receive the research results so that they can try to better understand the research results and discuss more concretely. The practitioners did not get enough time to think and discuss the research results as they were presented in Phase 2 of the interview which was conducted on the same day as Phase 1 was conducted. The practitioners were asked to give their opinion immediately after knowing the research results. Getting prior access to the research results will give practitioners time to think and reflect on it. We agree that allowing time between collecting prior and posterior probability will help in capturing true opinions and reflections. Although, booking two different interviews might be a challenge. In particular, it might be difficult if the benefits of the participation are not convincing or clear to the practitioners.

*Participating researchers:* To facilitate good discussions, the researchers participating in the knowledge translation process should be well informed about the topic for knowledge translation. Thus, it may not be suitable to use students, unless being PhD students focusing on the area being translated to industry.

## V. Conclusions

We have reported the feasibility implementation and evaluation of the KT framework in software engineering. The KT framework was successfully implemented and was perceived to be useful by the practitioners who were part of the evaluation. The framework is modular and we have implemented one instance of the KT framework. We plan to conduct further evaluations of the KT framework. We will consider the improvement suggested by practitioners in the feasibility

| | Knowledge identification | Transfer medium | Contextualising evidence |
|---|---|---|---|
| | Case 1 | | |
| Feasibility evaluation of KT1.0 | Knowledge identified through snowball strategy | Researchers provided knowledge in an interview | Bayesian synthesis was used to contextualise knowledge |
| | Case 2 (Future work) | | |
| Description and evaluation of KT 2.0 with different approaches | Knowledge identified through a mapping study | Use evidence briefings to transfer knowledge | Use Bayesian synthesis to contextualise knowledge |
| | Case 3 (Future work) | | |
| Description and evaluation of KT 2.0 with different approaches | Knowledge identified through a rapid review | Use evidence briefings to transfer knowledge | Use Bayesian synthesis to contextualise knowledge |

Fig. 6. Evaluation plan for future work

evaluation to improve implementation of the KT framework. Furthermore, we will implement the KT framework using different approaches to conduct the KT framework steps as shown in Figure 6. For example, we will use a transfer medium (evidence briefings [9]) for providing knowledge in a systematic way (part of Step 4), and rapid reviews [10] for identifying and selecting knowledge (KT Step 2).

## References


[1] C. L. Goues, C. Jaspan, I. Ozkaya, M. Shaw, and K. T. Stolee, "Bridging the gap: From research to practical advice," *IEEE Software*, vol. 35, no. 5, pp. 50–57, Sep. 2018.

[2] P. Devanbu, T. Zimmermann, and C. Bird, "Belief & evidence in empirical software engineering," in *Proceedings of the 38th international conference on software engineering*. ACM, 2016, pp. 108–119.

[3] D. Budgen, B. Kitchenham, and P. Brereton, "The case for knowledge translation," in *ACM / IEEE International Symposium on Empirical Software Engineering and Measurement*, Oct 2013, pp. 263–266.

[4] D. Badampudi, C. Wohlin, and T. Gorschek, "Contextualizing research evidence through knowledge translation in software engineering," in *Proceedings of the Evaluation and Assessment on Software Engineering*, ser. EASE '19. New York, NY, USA: ACM, 2019, pp. 306–311. [Online]. Available: http://doi.acm.org.miman.bib.bth.se/10.1145/3319008.3319358

[5] D. Badampudi and C. Wohlin, "Bayesian synthesis for knowledge translation in software engineering: Method and illustration," in *42th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*. IEEE, 2016, pp. 148–156.

[6] I. D. Graham, J. Logan, M. B. Harrison, S. E. Straus, J. Tetroe, W. Caswell, and N. Robinson, "Lost in knowledge translation: time for a map?" *Journal of continuing education in the health professions*, vol. 26, no. 1, pp. 13–24, 2006.

[7] R. K. Yin, *Case study research: Design and methods*. Sage publications, 2013.

[8] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in *Proceedings of the 18th international conference on evaluation and assessment in software engineering*. ACM, 2014, p. 38.

[9] B. Cartaxo, G. Pinto, E. Vieira, and S. Soares, "Evidence briefings: Towards a medium to transfer knowledge from systematic reviews to practitioners," in *Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*. ACM, 2016, p. 57.

[10] B. Cartaxo, G. Pinto, and S. Soares, "The role of rapid reviews in supporting decision-making in software engineering practice," in *Proceedings of the 22nd International Conference on Evaluation and Assessment in Software Engineering*. New York, USA: ACM, 2018, pp. 24–34.