A Framework to Validate Requirements Engineering Research Artefacts

Ban Al-Ani,

Dept. of Software Engineering, Faculty of Information Technology, University of Technology Sydney

alani@it.uts.edu.au

Abstract. Practitioners often resist the diverse assortment of research artefacts that are put forward to overcome the problems they face when developing requirements. This resistance often leads to the under-utilisation of artefacts despite their ability to reduce cost through software process improvement. There is a waste of resources (e.g. time and money) during development either as a consequence of practitioners' resistance or the inability of these artefacts to demonstrably meet practitioners' needs.

This paper details a three-layered evaluation framework consisting of several sets of criteria, standards and guidelines. The purpose of the framework is to provide a means to validate the diverse research artefacts produced thereby increasing the likelihood of its uptake by practitioners and academics alike. The paper also outlines ongoing work to implement the proposed framework to validate an analysis approach developed to assist stakeholders during the early stages of requirements development.

1 Introduction

The Requirements Engineering (RE) process is recognised as being the most critical process of software development. Errors made during this process can have disastrous effects on subsequent development steps, and on the quality of the resulting software.

Research into the RE process has produced an extensive body of knowledge and typically different types of methods, notations, or automated tools. The different products of research will be referred to collectively as *artefacts* herein. The term used is based on the Oxford English Reference Dictionary's definition of *artefact* (also artifact), which states that it is *a product of human art and workmanship* [18]. The term is used in this paper to refer to a product of academic research, which can be a framework, a set of guidelines, a model, a notation, a method, a tool...etc.

While some artefacts are successfully implemented in an academic environment few achieve widespread acceptance by practitioners (e.g. [6], [17]). This has been attributed to one or more factors, namely: the need for training that practitioners are unwilling to invest in, the artefact is unable to process non-textual requirements, or the performance degenerates when executed on different genres of text. Practitioners often consider research artefacts infeasible for these factors and others [10]. As a result, practitioners often resist adopting artefacts that support requirements activities despite their ability to reduce cost through software process improvement [13].

The purpose of this paper is not to inflict pessimism with regards to research conducted to support requirements development but rather to explore a means to provide a more optimistic future for the artefacts of research by making them more useful or 'marketable' to practitioners. One way to achieve this objective is to adopt the Research Artefact Validation (RAV) framework proposed in this paper.

The proposed framework is presented in the context of work within the field of benchmarking (e.g. [16]). It is currently made up of a combination of existing criteria, categorised into steps and grouped within one of three layers. The purpose of the framework is to assist researchers establish that a method, for example, accomplishes what practitioners really require. The framework also provides a means to formalise an artefact's validation process. A method that meets the criteria within RAV will be certified and thus provides researchers with a means to guarantee that the artefact was subjected to a formal validation process.

The paper is made up of two main sections. The first section presents details of the framework, whereas the second section reports ongoing work to implement the framework to validate an analysis approach. Finally, the paper concludes with remarks about the framework and an outline of the tasks that are yet to be accomplished.

2 Definition of the Research Artefacts' Validation (RAV) Framework

Sim et al [16] state that there is a need to identify benchmarks within the software engineering community to identify the key problem areas and encapsulate this knowledge in an evaluation. They set out to define benchmarks as tests or sets of tests used to compare the performance of alternative tools or techniques. RAV research does not limit itself to tools or techniques alone but also sets out to validate models, notations and other artefacts produced as a result of research. However, RAV research does focus itself on the performance measure component defined in the Sim et al work, in that RAV does set out to validate artefact performance, which is defined as the fitness to purpose. The framework can be considered a benchmark that operationalises the concept of validation.

The RAV framework is based on the premise that the practitioners are the customers for academics working in the field of research and development. Consequently, the needs of practitioners are paramount when developing methods, notations or automated tools. These needs in addition to the benchmarks set by the academic community and established formal standards dictate the nature of research conducted in a particular area and the characteristics of any artefacts developed during or as a result of academic research projects. Consequently, the proposed framework consists of three layers, namely: the *standards layer*, the *empirical layer*, and the *industry layer*. Each layer is made up of one or more *steps* to indicate that it is part of the validation process within the defined framework. A more detailed definition of the three layers and the steps within each of these layers is presented in the following sections.

2.1 Standards Layer

The layer consists of a collection of standards that are developed by official-standards-making groups. Each set of standards is regarded as a step within the layer. Some of the groups that impact computer standards are listed below:

- International Standards Organisation (ISO).
- International Telecommunications Union (ITU).
- National standards e.g. American National Standards Institute (ANSI).
- Institute of Electrical and Electronics Engineers (IEEE).

These groups develop formal standards for the products and process' of systems development, which includes those carried out during requirements engineering. Two examples of IEEE Standards concerned with requirements are IEEE Standards 830 (Software Requirements Specification) and IEEE Task Force on Engineering of Computer-Based Systems (ECBS)-Requirements. Each of these standards represents a *step* within the standards layer.

While government bodies charter standards with strict procedures and rules imposed on the standards development process, the fact remains that there remains little or no communication between the developers of the standards and the engineers who implement those standards [14]. Therefore an artefact's conformance to well-defined formal standards does not guarantee that it is practical to implement it. Researchers cannot rely on an artefact's conformance to formal standards as means to validate it within the real world. This suggested the need for an additional category and consequently layer of validation.

2.2 Empirical Layer

The layer consists of a collection of empirical methods designed to validate a particular artefact. Each empirical method is regarded as a step within the layer. In addition to these well-defined empirical methods there are also methodologies that can assist researchers choose which to adopt.

Kitchenham et al [9] identify nine methods of evaluation and a set of criteria to help evaluators select an appropriate method. The methodology, DESMET, is intended to support individuals performing the following roles:

- A vendor of a method or tool seeking to demonstrate the advantages of a product.
- A software engineer responsible for assessing a proposed process change.
- An academic researcher developing or investigating a new method

DESMET is one of many existing methods (e.g. [5]) proposed to assist researchers decide on the most appropriate empirical study to adopt during validation. These can also be included as individual steps within this layer as an alternative to or in conjunction with DESMET.

Researchers generally consider validation through empirical studies a natural part of the research and often do not rely on empirical studies alone. The need for additional validation remains because while empirical studies can suggest certain conclusions there exists a risk that what is concealed is vital.

2.3 Industry Layer

The layer consists of a collection of standards that are defined by practitioners. Each set of standards is regarded as a step within the layer.

Several surveys have been conducted in an attempt to identify problems with current RE practices and develop solutions to support and improve current practices (e.g. [4], [7]). However, despite the successful identification of problems and solutions, a more recent survey conducted by Juristo et al [8] provides evidence that this gap still exists. One recommendation put forward by Juristo et al to researchers is that they provide more elaborate and practical guidelines for transferring recommendations to industry.

Workshops have also been conducted in an attempt to investigate problems associated with RE research, development and industrial uptake [11]. The objective of the workshop was to generate a report (DGIII) that can be used by a sector of the European Commission (EC) to assess research. The list of recommendations and criteria proposed by Morris, et al [11] can serve as the initial step within this layer.

A review of literature revealed that industry-defined standards or criteria is not extensively documented or reported. The Morris et al [11] report is the only report found which states that a consensus was attained from a large set of representatives from different industrial and academic sectors across Europe on the issues concerning industrial uptake of research and development projects in RE. This leads to the conclusion that this area needs to be investigated further. There also exists a need to establish whether these issues are shared by industrial and academic sectors of other regions and cultures.

An overview of the proposed RAV framework layers and some of the steps identified thus far is presented in figure 1. The figure demonstrates that there is no set sequence to adopting one of the proposed layers as a means of validation. It is only for explanatory purposes that the framework is presented in this sequence. The proposed framework steps can also be conducted concurrently.

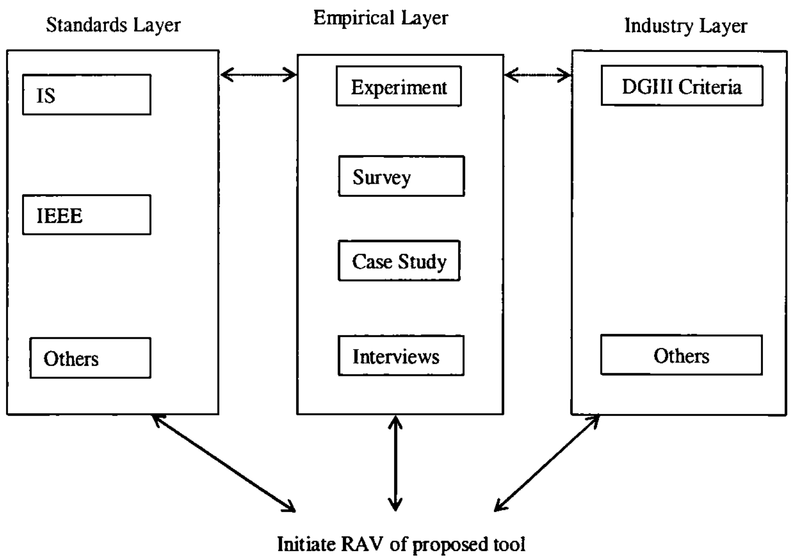

Fig. 1. An abstract representation of the proposed RAV Framework: The layers are represented as boxes, which contain pre-defined steps. The lines between layers represent the validation path that an artefact might take when being validated. An artefact that is successfully subjected to this process would attain the RAV Framework seal of success. The seal would indicate that the artefact achieved the level of maturity necessary to make it viable for use by practitioners

An artefact that is successfully subjected to this process would attain the RAV Framework *seal of* success. The seal would indicate that the artefact achieved the level of maturity necessary to make it viable for use by practitioners. To this end the framework must gain recognition by an official body that is an accepted authority in both the academic and practitioners' communities. An effort must be made to find an appropriate organisation that is not only willing to adopt such a framework but capable of credibly implementing it.

3 Case Study: Applying the RAV Framework to QSARA

The researcher is currently applying the RAV framework to validate the Qualitative Systematic Approach to Requirements Analysis (QSARA), which was developed to assist stakeholders during the early stages of requirements development [2].

Steps within the empirical layer were successfully implemented to validate the approach and work in the industry layer is currently in progress. An outline of steps within the RAV framework layers that were implemented are presented in the following sections.

3.1 Standards Layer

The researcher reviewed relevant standards to verify the approach's compliance and to prepare for its formal assessment by assessors. The ISO standards ISO/IEC 15504 concerned with process improvement, software requirements analysis and software engineering process assessment are some of the standards relevant to the analysis approach. Further work is necessary to formally validate the approach's compliance to these standards.

3.2 Empirical Layer

Extensive work was conducted to validate QSARA through empirical research methods. A combination of these methods was adopted, namely:

1. The researcher collaborated to conduct an empirical study in which seventy participants took part. The study consisted of two components: experiment and survey.

   The chosen studies were conducted correctly based on the principles stated by Yin, 1994 and Kitchenham et al, 1994. The survey questionnaire was developed based on guidelines defined by Seymour [15].

   Participants were randomly allocated to one of three groups, each of which would have well-defined tasks, namely:

   Group1 used an analysis method they are familiar with to analyse a given document using the tool(s) they consider necessary to capture and maintain their analysis.

   Group 2 used the proposed QSARA to analyse the requirements document and use the support tool(s) they consider useful in capturing and maintaining their analysis.

   Group 3 used QSARA to analyse the document and QuestMap to capture and maintain their analysis.

   All three groups were presented with the same draft of a requirements document and were asked to construct a description of the same feature captured within the document. All three groups were given the same amount of time to conduct the tasks.

   Data gathered at the end of this empirical study supported that the proposed QSARA did assist analysts construct more comprehensive descriptions and did assist analysts in detecting inter-dependency relationships between features.

2. The researcher conducted a study in a more controlled experimental environment. Forty-five participants took part in the experiment and responded to a pre-experiment and post-experiment questionnaire. The experiment was conducted over a three-day period to demonstrate that results are repeatable.

   Data collected during this second empirical study is currently being analysed.

A more detailed report of the empirical research will be published at a future date.

The empirical work findings statistically supported research hypotheses. However, the researcher is a firm believer of the saying (attributed to Benjamin Disraeli) that there are *'lies, damned lies, and statistics'*. Hence, further validation was sought in another of the RAV framework layers- the industry layer.

3.3 Industry Layer

Research is ongoing to determine national and international criteria that can be used to validate QSARA. At present, the approach is based on the findings of earlier work [1].

A link with local organisation is established and practitioners will validate the approach based on the standards of their particular organisation's criteria. The initial set of criteria was outlined during the first few face-to-face meetings. Further details will be collected through text and report analysis in addition to conducting a survey.

The plan is to use the QSARA to analyse a comprehensive tool for sustainable building design and construction within an organization. The team responsible for developing the tool recognised the need to make it available online for use by developers as part of the development application process. Initially, the online tool will be trialled with seven other organisations in NSW. However, further QSARA development and validation is necessary to establish that it is a suitable approach to use by distributed groups. A plan to initiate research into this area is underway.

Several working groups are involved in developing the online tool and they face problems often associated with requirements management and group management, in addition to the problems that often arise in requirements engineering. Implementing the proposed approach within this environment is still in its infancy and conclusions are yet to be drawn from the results.

4 Concluding Remarks

Despite the great leaps and jumps forward in research and development, practitioners have survived attempts to overcome some of the imperfections of RE practices. The paper presented a validation framework consisting of three layers. The researcher sought to utilise existing guidelines, standards and criteria by incorporating them as steps within the layers. The framework is by no means complete and work is ongoing to elaborate the layers further by adding more steps in each, using published work rather than starting from scratch and reinventing the wheel. The objective is to construct a framework that validates the diverse artefacts produced. The complete framework would then be subjected to self-validation in addition to validating RE artefacts widely accepted and utilised in industry. These projects must be put on hold until the industry standards layer is well-defined because very little work focuses on establishing widely accepted industry criteria.

Adopting the proposed framework can lead to the following:

1. Improve the quality of a research artefact by submitting it to various categories of validation thereby increasing the likelihood of identifying possible flaws.
2. Increase the credibility of an artefact, which in turn can increase the likelihood of its uptake by practitioners and academics alike

A more mature framework would need to be adopted by an official body that is an accepted authority in both the academic and practitioner's communities. A RAV seal of success would then be a credible indication that the artefact's is both practical and feasible. It would also indicate that the artefact has been validated through one or more empirical methods, formal and industrial standards in accordance to established guidelines.

The paper also presented an implementation of the proposed framework on an analysis approach. Steps in the empirical layer and industry layer were outlined. Details of its implementations and results achieved will be published at a future date. Further work is necessary to validate the approach's adherence to formal standards.

The framework is still in a state of evolution and the process of identifying potential steps within each layer is ongoing. Research will also be conducted to investigate the possibility of developing the framework further to include steps that will assist researchers and practitioners validate artefacts intended to support other software development activities e.g. design, programming. The feasibility of such a framework would need to be considered in addition to its practicality.

References

1. Al-Ani, B., (2001). "A Framework for the Detection and Evaluation of Partially Complete Requirements Documents", Doctor of Philosophy Dissertation in Computer Systems Engineering, Faculty of Engineering, University of Technology Sydney.
2. Al-Ani, B., (2002). "A Qualitative Systematic Approach to Requirements Analysis (QSARA)", International Conference on Software Engineering Research and Practices, Las Vegas, pp. 385 - 391.
3. Christensen, Larry B., (2001). Experimental Design, Eighth Edition, Allyn and Bacon.
4. Curtis, B., Krasner, H., Iscoa, N., (1988). "A Field Study of the Software Design Process for Large Systems", Communications of the ACM, Vol. 31, No 11, Nov. 1988, 1268- 1287.
5. Donohue, S., and Dugan, J., (2003). "Modelling the good enough to release Decision using V & V Preference Structures and Bayesian Belief Network", Proceedings of the Annual Reliability and Maintainability Symposium, pp. 568-573.
6. Edwards, M. L., Flanzer, M., Terry, M., Landa, J., (1995). "RECAP: A Requirements Elicitation, Capture, and Analysis Process Tool for Large Complex Systems", Proceedings of the 1st International Conference on Engineering of Complex Computer Systems, Institute of Electrical and Electronics Engineers, Aug., pp. 278-281.
7. El Emam, K., Madhavaji, N., (1995). "A Field Study of Requirements Engineering Practices in Information Systems Development", IEEE 2nd Symposium for Requirements Engineering, CS Press, Loss Almitos, California, July, pp. 68-80.

8. Juristo, N., Moreno, A., Silva, A., (2002). "Is the European Industry Moving Towards Solving Requirements Engineering Problems?", IEEE Software, November-December, pp. 70-76.

9. Kitchenham, B., Linkman, S., Law, D., (1997). "DESMET: A Methodology for Evaluating Software Engineering Methods and Tools", Computer and Control Engineering Journal, June, pp. 120-126.

10. Mead, N. (2000). "Why is it so difficult to introduce Requirements Engineering Research Results into Mainstream Requirements Engineering Practice?", 4th International Conference on Requirements Engineering (ICRE'00), pp. 75.

11. Morris, P., Masera, Wilikins, (1995). Panel on Industrial Workshop on Requirements Engineering,, 3rd IEEE International Symposium on Requirements Engineering (RE'97) , January 05 - 08, pp. 41.

12. Recommended Practice for Software Requirements Specifications (ANSI/IEEE), October 1998. ISBN 1-7381-0332-2.

13. Sawyer, P., Sommerville, I., Viller, S., (1999). "Capturing the Benefits of Requirements Engineering", IEEE Software, March/April, pp. 78-85.

14. Severance, C., (1995). "The Value of the Formal Standards Process", Computer, August, pp. 82-83.

15. Seymour, S., Norman M.B., Norbert, S., (1996). Thinking about answers: the application of cognitive processes to survey methodology. San Francisco, Jossey-Bass Publishers.

16. Sim, S., Easterbrook, S., Holt, R., (2003). "Using Benchmarking to Advance Research: A Challenge to Software Engineering", IEEE Proceedings of the 25th International Conference on Software Engineering (ICSE'03).

17. Smith, T. J., (1992). "READS: A Requirements Engineering Tool", IEEE CS Press, Loss Almitos, California, Jan., pp. 94-281.

18. Oxford English Reference Dictionary (1996). Oxford University Press

CERE'03

Proceedings of the 1st International Workshop on Comparative Evaluation in Requirements Engineering (CERE'03)

Edited by
Vincenzo Gervasi, Didar Zowghi, Steve Easterbrook, Susan Elliott Sim

Held in conjunction with the 11th IEEE International Requirements Engineering Conference (RE'03)
8-12 September 2003, Monterey Bay, California, USA

Published by
Faculty of Information Technology
University of Technology, Sydney
PO Box 123
Broadway NSW 2007

ISBN: 1-86365-861-0

Table of Contents

Workshop Summary by Vincenzo Gervasi, Didar Zowghi,
Steve Easterbrook and Susan Elliott Sim....................v

Workshop Organisation....................ix

Session 1. Positions....................1

Methodologies of Requirements Engineering Research and Practice: Position Statement....................3
Wieringa, Roel J.

Comparative Evaluation: Implications from the Multidisciplinary Nature of Requirements....................9
Wasson, Kimberly S.

Session 2. Frameworks....................15

The Case for Applying an Early-Lifecycle Technology Evaluation Methodology to Comparative Evaluation of Requirements Engineering Research....................17
Feather, Martin S.

A Framework to Validate Requirements Engineering Research Artefacts....................27
Al-Ani, Ban

Session 3. Experiences....................35

A research methodology in the study of requirements negotiations in geographically distributed software teams....................37
Damian, Daniela E.

Experiences from Lightweight RE Method Evaluations....................49
Nikula, Uolevi