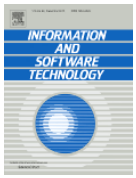

Information and Software Technology
ISSN: 0950-5849, Impact Factor: 2.694, 5-Year Impact Factor: 2.924

Call for Papers

Special Issue on Enhancing Credibility of Empirical Software Engineering

AIM AND SCOPE

Researchers continuously struggle to provide sufficient evidence regarding the credibility of their findings. At the same time, practitioners have difficulties in trusting the results with limited credibility. Probably the most striking summary of the research crisis in multiple disciplines is given by Ioannidis who in his seminal paper [3] (with more than 4000 citations) claims that "*Most Research Findings Are False for Most Research Designs and for Most Fields*". According to Gartner, the size of the worldwide software industry in 2013 was US$407.3 billion [1]. Hence, invalid recommendations or missing research findings in software engineering can cost a lot of money.

Problems with the credibility of research findings are not absent in software engineering as well. For example, Shepperd et al. [8] meta-analysed 600 experimental results drawn from primary studies that compared methods for predicting fault-proneness. They found that the explanatory factor that accounted for the largest percentage of the differences among studies (i.e., 30%) was research group. In contrast prediction method, which was the main topic of research, accounted for only 1.3% of the variation among studies. Hence, they commented that there seems little point in conducting further primary studies until the problem that "*it matters more who does the work than what is done*" can be satisfactorily addressed.

This special issue focuses on the two complementary and important areas in software engineering research: 1) **reproducible research**, and 2) **modern statistical methods**. Reproducible research refers to the idea that the ultimate product of research is the paper plus its computational environment. That is, a reproducible research document incorporates the textual body of the paper plus the data used by the study, and the analysis steps (algorithms) used to process the data. The reason for adding the whole computational environment is that other researchers then can repeat the studies and reproduce the results, which in turn would deliver more credible (trustworthy) results. Reproducibility is a crucial aspect of credible research. Unfortunately, it is often impossible to reproduce data analyses, due to lack of raw data, sufficient summary statistics, or undefined analysis procedures. Thus wider adoption of reproducible research would be beneficial for Empirical Software Engineering [7].

Furthermore, true research findings may be missed due to inadequate statistical methods that do not reflect the state of the art in statistics, when modern statistical methods including robust [4], [5], Bayesian [2] and meta-analysis methods [6] are available. Statistical techniques widely used in Empirical Software Engineering studies base, to a large extent, on two fundamental assumptions: normality and homogeneity of variances. These techniques are often considered robust when either of these assumptions is violated. Unfortunately, recent research findings provide evidence that widely used classic methods can be highly unsatisfactory for comparing groups and studying associations [9]. A fundamental problem is that violating the basic assumptions underlying statistical methods can result in relatively low power or missing important features of the data that have practical significance. Low power not only increases the false negatives probability meaning that potentially valuable discoveries may be lost, but also leads to inflated effect sizes for true positives, which can lead to under-powered replications and failure to confirm true results. In addition, Null Hypothesis Statistical Testing (NHST) and p-values remain the standard inferential tool in many disciplines including software engineering, in spite of the availability of alternative more trustworthy approaches, e.g., inference based on confidence intervals (CIs) instead of p-values or Bayesian approaches to avoid the pitfalls of NHST.

The aim of the special issue is to stimulate awareness and uptake of recent advances in these crucial areas, reproducible research and modern statistical methods, by the software engineering community. That is to increase the uptake of reproducible research methods and tools, as well as robust, Bayesian and meta-analysis statistical methods. In particular the objective is show examples of empirical software engineering research which employ the aforementioned methods and tools to evaluate software engineering methodologies, practices, technologies and tools to allow more credible evidence-based decisions.

Call for Papers is available at http://madeyski.e-informatyka.pl/download/ISTSpecialIssue.pdf and
http://www.journals.elsevier.com/information-and-software-technology/call-for-papers/special-issue-on-enhancing-credibility-of-empirical-software

Topics of Interest

We solicit high quality research articles, guidelines and review articles concerning quantitative and/or qualitative empirical software engineering research and practice focused on topics which include, but are not limited to, proposals, uses, reviews and/or evaluations of:

- Reproducible research tools or methods (e.g., employing reproducible research in empirical software engineering).
- Statistical methods addressing the pitfalls of the classic statistical methods or leading to more trustworthy results (e.g., employing robust statistical methods, Bayesian methods or meta-analyses in empirical software engineering).

Using both, a reproducible research environment and modern statistical methods (e.g., robust methods, Bayesian methods, meta-analyses) in software engineering (e.g., to perform empirical evaluation of methods, practices, methodologies, technologies and tools in software engineering), would be an additional advantage. Other means to enhance credibility of empirical software engineering are also within the scope of the special issue.

Submission and Important Dates

The special issue's paper submission page is available at https://www.evise.com/profile/#/INFSOF/login. When submitting the manuscript for this special issue, please select "*Special issue: Enhancing Credibility of Empirical SE*" as the issue type. Formatting templates can be found at https://www.elsevier.com/authors/author-schemas/latex-instructions Please note that the journal of *Information and Software Technology* prescribes the use of a "*structured abstract*" including the following components: Context, Objective, Method, Results and Conclusions. All contributions must not have been previously published or be under consideration for publication elsewhere. A submission extended from a previous conference version must have at least 30% new material.

Tentative Timeline:

Submission deadline:	May 31, 2017
Notification of first decision:	July 16, 2017
Submission of revised manuscripts:	September 4, 2017
Re-reviews completed:	October 27, 2017
Minor revision due:	November 24, 2017
Tentative acceptance deadline:	December 9, 2017

Guest Editors of the Special Issue:

Lech Madeyski
Wroclaw University of Science and Technology
Poland

Barbara Kitchenham
Keele University
UK

Krzysztof Wnuk
Blekinge Institute of Technology
Sweden

Email to Guest Editors:
mailto:IST.SI@pwr.edu.pl

Call for Papers:
http://madeyski.e-informatyka.pl/download/ISTSpecialIssue.pdf

http://www.journals.elsevier.com/information-and-software-technology/call-for-papers/special-issue-on-enhancing-credibility-of-empirical-software

References

[1] Gartner says worldwide software market grew 4.8 percent in 2013. Available at http://www.gartner.com/newsroom/id/2696317.

[2] Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. *Bayesian Data Analysis*. CRC Press, 2013.

[3] John P. A. Ioannidis. Why Most Published Research Findings Are False. *PLoS Medicine*, 2(8):696–701, 2005.

[4] Barbara Kitchenham. Robust Statistical Methods: Why, What and How: Keynote. In *Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering (EASE 2015)*, pages 1:1–1:6, 2015. doi:10.1145/2745802.2747956.

[5] Barbara A. Kitchenham, Lech Madeyski, David Budgen, Jacky Keung, Pearl Brereton, Stuart Charters, Shirley Gibbs, and Amnart Pohthong. Robust Statistical Methods for Empirical Software Engineering. *Empirical Software Engineering*, 22(2):579–630, 2017. URL: http://dx.doi.org/10.1007/s10664-016-9437-5, doi:10.1007/s10664-016-9437-5.

[6] Barbara Ann Kitchenham and Lech Madeyski. Meta-analysis. In Barbara Ann Kitchenham, David Budgen, and Pearl Brereton, editors, *Evidence-Based Software Engineering and Systematic Reviews*, chapter 11, pages 133–154. CRC Press, 2016. Available at https://www.crcpress.com/Evidence-Based-Software-Engineering-and-Systematic-Reviews/Kitchenham-Budgen-Brereton/p/book/9781482228656.

[7] Lech Madeyski and Barbara Kitchenham. Would wider adoption of reproducible research be beneficial for empirical software engineering research? *Journal of Intelligent & Fuzzy Systems*, 32:1509–1521, 2017. URL: http://madeyski.e-informatyka.pl/download/MadeyskiKitchenhamJIFS.pdf, doi:10.3233/JIFS-169146.

[8] Martin Shepperd, David Bowes, and Tracy Hall. Researcher Bias: The Use of Machine Learning in Software Defect Prediction. *IEEE Transactions in Software Engineering*, 40(6):603–616, 2014. doi:10.1109/TSE.2014.2322358.

[9] Rand R. Wilcox. *Introduction to Robust Estimation and Hypothesis Testing*. Elsevier, 3rd edition, 2012.