Reviewing the Systematic Review

William Wallis — 2025138W

December 19, 2016

Abstract

Systematic reviewing is a technique for bringing scientific rigour to a computer science literature review, pioneered by Barbara Kitchenham (Kitchenham, 2004). Specifically, Kitchenham's systematic reviews utilise concepts from the field of medical research to create literature reviews which are repeatable, and produce statistical and empirical results. 12 years after Kitchenham's original guidelines were set for structuring a systematic literature review, the technique has seen widespread adoption — but the original guidelines raise questions and note possible issues with the method. With a wide set of samples to choose from, a review of these systematic reviews may highlight whether these concerns are worth revisiting, before Kitchenham's guidelines — or other methods derived from them — become standard practice for the software engineering research community.

1 Introduction

A Systematic Review is a literature review which collates the results of many papers, using statistical analysis to draw empirical results about the state of a research field and to answer research questions posited as the motivation for the review. Systematic reviews are born from the philosophy that a literature review should have scientific merit, and produce reproducible results, rather than standard techniques for literature reviews, which leave more room for subjective insight. The scientific nature of a systematic literature review, in theory, removes ambiguity and bias from literature review practice and lends the review the same validity and credence as a research study.

Kitchenham's systematic literature review technique has begun to dominate as a literature review technique for software engineering. Kitchenham's review procedure stems largely from literature review techniques in medical research (Kitchenham, 2004; Khan et al., 2001), where empirical studies which verify the validity of literature already published in peer-reviewed journals is paramount to civilian safety. While the technique has seen widespread adoption, some issues exist with the implementation — as noted by Kitchenham herself in a systematic review of systematic review procedures (Kitchenham

and Brereton, 2013). More fundamentally, in Kitchenham's original guidelines there exist some notes which cast doubt on the suitability of a systematic review in the field of software engineering. For example:

> In particular, software engineering research has relatively little empirical research compared with the large quantities of research available on medical issues, and research methods used by software engineers are not as rigorous as those used by medical researchers.

In her guidelines, Kitchenham provides types of empirical data which software engineering research *does* produce which can be appropriate for analysis in a literature review. However, whether research rigour and types of data collected are make appropriate note of by the research community is clear only now that a wealth of systematic literature reviews have been produced.

In this review, a series of systematic literature reviews will be analysed and searched for their scrutiny of research rigour and format of empirical data. In this way, the importance of this doubt regarding the suitability of systematic reviews for software engineering research will be assessed.

References

Khalid S Khan, Gerben Ter Riet, Julie Glanville, Amanda J Sowden, Jos Kleijnen, et al. *Undertaking systematic reviews of research on effectiveness: CRD's guidance for carrying out or commissioning reviews.* Number 4 (2n. NHS Centre for Reviews and Dissemination, 2001.

Barbara Kitchenham. Procedures for Performing Systematic Reviews. 2004.

Barbara Kitchenham and Pearl Brereton. A systematic review of systematic review process research in software engineering. *Information and Software Technology*, 55 (12):2049–2075, 2013. ISSN 09505849. doi: 10.1016/j.infsof.2013.07.010. URL http://dx.doi.org/10.1016/j.infsof.2013.07.010.