Protocol for the Systematic Literature Review on Web Development Resource Estimation

Author: Damir Azhar

Supervisor: Associate Professor Emilia Mendes

Table of Contents

List of Tables

1. Background

Effective resource management is crucial for successful software development. It enables managerial decisions to be made with regards to cost, quality, and scheduling tradeoffs inherent in the software lifecycle [1, 2]. Mendes described resources as being any "factors such as cost, effort, quality, 'problem size' that have a bearing on a project's outcome" [3]. Early resource estimation was centered on the use of a size measure (e.g. lines of code or function points) as a key determinant to software development effort, and hence resource requirements. However using problem size as the main predictor of resource requirements provides only limited support for managerial decision making [1]. A more comprehensive approach would involve expanding the focus to include other factors relevant to software development, as well as taking into account the causal relationships between them.

In their 2004 paper [1], Fenton et al. described using a Bayesian network to form a causal, non-deterministic model for making resource decisions for general software development. Their model considered resources associated with various aspects of the development process, including project requirements and specifications, the quality of the development team involved, the effort required to implement the necessary functionality, and the quality of the final product. The resulting decision support tool based on this model is part of the MODIST toolset that deals specifically with "risk management for distributed software development" [1]. MODIST has since been incorporated into the AgenaRisk software system [4]; a commercial software system that uses Bayesian networks to provide risk analysis and decision support for a variety of industry sectors (e.g. banking, aerospace, and defense industries).

Web development is a relatively new and rapidly growing industry, with the number of Web development companies in the United States alone increasing from less than 1000 businesses in 1995 to over 30,000 in 2005 [5]. By 2010 the Web development industry is expected to experience a further growth of over 20%. This would make research geared towards enabling Web development companies to make more efficient managerial decisions very worthwhile. Simply porting over existing software resource estimation methodologies would not be adequate as Web development is different from general software development [6]. Web applications tend to vary widely in terms of structure and implementation from project to project, often make use of numerous non-code elements (e.g. multimedia objects) and may have to work with different (and possibly incongruous) legacy systems [6].

Considering the importance that Web development plays in today's industry and its difference from general software development, a systematic literature review would be essential in establishing the current state of the art as well as document existing gaps in the field of Web resource estimation. This systematic review would thus be geared at "identifying, evaluating and interpreting all available research" [7] relevant to resource estimation for Web development. The systematic literature review process can be divided into three phases namely planning, conducting and reporting the review [7]. The following document specifies the protocol that the review process will follow and hence represents the first phase of this process.

2. Research Questions

Formulating the research questions that a systematic literature review will address is central to the review process [7]. The research questions determine which primary studies are selected, the data to be extracted from these selected studies, and how this data is to be analyzed so that the research questions can be answered. One approach to formulating research questions is to use the PICOC criteria specified by Petticrew and Roberts [8], which structures research questions according to five attributes: population, intervention, comparison, outcome and context. However, since the focus of this literature review is not to compare interventions, the comparison attribute will not be utilized and hence only the population, intervention, outcome and context (PIOC) attributes of the research questions are shown in Table 1.

Population	Web development projects
Intervention	Methods/techniques used for Web resource estimation, resource predictors considered, characteristics of the datasets worked on
Outcome	Accuracy of the methods/techniques used for Web resource estimation
Context	Within the domain of Web development with a focus on empirical studies

Table 1: Research questions as structured by the PIOC criteria.

Therefore in order to identify and evaluate all the research done on Web resource estimation, the research questions that need to be addressed by our systematic literature review are as follows:

Question 1:
What methods and techniques have been used for Web resource estimation?

Question 1a:
What metrics have been used to measure estimation accuracy?

Question 1b:
What (numerical) accuracy did these various methods/techniques achieve?

Question 2:
What factors (e.g. effort, quality, size) have been investigated as resource predictors for Web development?

Question 2a:
What resource predictors have been used in the estimation process?

Question 2b

At what stage are these resource predictors gathered?

Question 3:

What are the characteristics (single/cross-company, student/commercial projects) of the datasets used for Web resource estimation?

3. Identifying Relevant Literature

The purpose of a systematic literature review is to comprehensively identify all the studies that are relevant to the research questions the review is to address [7]. Not only does the identification process need to be rigorous, it also needs to be unbiased. In order to minimize researcher bias a pre-defined strategy for this identification process is required, and will be described in the following section.

3.1. Deriving the Search Terms

The first step in finding studies relevant to the research questions that need to be addressed, is to identify search terms that will be used in the search process. These search terms can be regarded as the key elements that underlie the research questions and can be identified as follows [7]:

- The four PIOC facets (population, intervention, outcome and context) can be viewed as the key elements underlying the research questions. Table 2 contains the search terms derived from the research questions structured according to the PIOC criteria (as specified in Table 1).
- Subject headings/keywords used by related articles in journals and databases can also be used as a source of search terms (see Table 3).
- Synonyms, alternate spellings, and abbreviations of the search terms derived by the previous two steps should be considered (see Table 4).

Once the search terms have been identified, they can be compiled into a search string that will be used in the search process. This can be done using the Boolean operators OR and AND as follows:

- The OR operator will be used to group the various forms (e.g. synonyms and alternate spellings) of individual search terms (see Table 5).
- The AND operator will then be used to link the different search terms into a single search string (see Table 6).

Population	Web development, projects
Intervention	Methods, techniques, Web resource estimation, predictors, datasets
Outcome	Accuracy, estimation, Web resource predictors
Context	Web development, empirical studies

Table 2: Search terms derived from PIOC.

Umbers and Miles (2004)	*Web applications*, design patterns, *effort estimation*, function point size measurement, *resource estimation*, software development time, software engineering, software project
Baresi and Morasca (2007)	W2000, *Web application* design, *effort estimation*, empirical study
Mendes and Mosley (2001)	*Web application* characteristics, Web page design, *World Wide Web* authoring, boxplots, case study evaluation, *effort prediction* models, linear regression, residuals, *resource management*, software project management, statistical techniques, stepwise multiple regression
Mendes (2000)	*World Wide Web applications*, authoring, development effort, *hypermedia processes*, quantitative case study, representational measurement theory, software development *effort prediction* model, software metrics, statistically significant correlations
Mendes et. al (2001)	*Web application characteristics* measurement, *Web development*, Web site design, *World Wide Web* authoring, algorithmic models, case study evaluation, flexible techniques, linear regression, *prediction* power, project lifecycle, *resource management*, software development organizations, software *effort estimation*, software engineering approach, software project management, stepwise multiple regression, testing
Tilley (2001)	*World Wide Web*-site log analyzer program modification, application domain novelty, case study, *net-centric computing applications*, software development *effort estimation*, student inexperience, undergraduate course offerings, undergraduate software engineering class
Morisio et. al (1999)	*Web based application* functionality, case study, classical function points, continuous calls, *cost estimation* model, development strategies, *effort prediction* model, lines of code, object oriented framework based development, object oriented function points, reuse types, size measurements

Table 3: Search terms (italicized) derived from keywords found in relevant papers.

Web	World Wide Web, Web based, Web-based, net-centric, hypermedia, Web hypermedia
Resource[1]	Cost, effort, maintenance, maintainability, quality, reliability
Estimation	Prediction, forecasting, calculation
Method	Process, technique, system, practice, procedure, model

Table 4: Search terms derived from Tables 2 and 3, along with their alternate spellings and synonyms.

Web OR net-centric OR hypermedia
resource OR cost OR effort OR maintenance OR maintainability OR quality OR reliability
estimation OR prediction OR forecasting OR calculation
method OR process OR technique OR system OR practice OR procedure OR model

Table 5: Concatenation of synonyms and alternate spellings using the Boolean OR operator.

(Web OR net-centric OR hypermedia) AND
(resource OR cost OR effort OR maintenance OR maintainability OR quality OR reliability) AND
(estimation OR prediction OR forecasting OR calculation) AND
(method OR process OR technique OR system OR practice OR procedure OR model)[2]

Table 6: Construction of search string by concatenating search terms using the Boolean operator AND.

3.2. The Search Process

With the search terms identified, the search process can now begin. As discussed at the beginning of this section, the aim of a systematic literature review is to comprehensively identify all the studies that are relevant to the research questions the review is to address. The search process therefore has to be a rigorous one, and can be split into a primary and secondary search phase as will be subsequently detailed.

3.2.1. Primary Search Phase

This phase involves identifying and searching through primary sources of relevant literature. These sources include online databases, search engines, conference proceedings and grey

---

[1] It needs to be noted that while cost and effort are often used synonymously, they are not synonyms of resource. However, both cost and effort estimation have been used as the basis for resource estimation as seen in Table 3.

[2] The last set of search terms was eventually removed as no results were being returned with its inclusion.

literature (e.g. PhD theses and technical reports). Given that resource estimation for Web development is the focus of this literature review, and that the World Wide Web started as a CERN project in 1989 with the first Web browser Mosaic appearing in 1993 [9], the primary search phase will only consider literature published from 1990 (inclusive) onwards. The list of primary sources is given below. These resources were recommended by the University of Auckland Library website as resources relevant to Computer Science, or by existing literature relating to systematic literature reviews in software engineering [7, 16].

Online Databases

- INSPEC
- IEEE Xplore
- ACM Digital Library
- ProQuest Computing
- Current Contents
- Web of Science
- Computer Database
- SCOPUS
- ScienceDirect
- Springer Link

Online Search Engines

- CiteSeerX (http://citeseerx.ist.psu.edu/)

PhD Theses

- ProQuest Theses and Dissertations

3.2.2. Secondary Search Phase

The secondary search phase completes the search procedure by ensuring that the primary search phase has not missed any relevant literature. The secondary search phase entails:

- Reviewing the reference lists of all the literature retrieved by the primary search phase. This procedure is iterative being repeated on any new literature found.
- Key researchers identified by the primary search phase are contacted with regards to any further relevant research, including unpublished papers and technical reports.

3.3. Search Process Documentation

In order for a systematic literature review to be regarded as reliable, the review process needs to be both transparent and (to a degree) repeatable [7]. This means that the search process has to

be carefully documented. We will use the documentation procedures outlined by Kitchenham in her guidelines for performing a literature review [7].

Data Source	Documentation
Online Databases/Digital Libraries	• Name of database • Search strategy for the database • Date of search • Years covered by search
Search Engines	• Name and URL of search engine • Search strategy for the search engine • Date of search
Journal Hand Searches	• Name of journal • Years searched • Issues not searched (if any)
Conference Proceedings	• Title of proceedings • Name of conference (if different) • Journal name (if published in a journal)
Unpublished Research	• Research group and researchers contacted (full contact details) • Research Website searched (date and URL)
Other Sources	• URL • Date searched/contacted

Table 7: Procedure for documenting the search process, based on those outlined in [7].

4. Study Selection Criteria

Once the primary studies have been identified, the next step in the systematic literature review process is to assess them in order to select those that best answer the research questions the review is to address. As with the search process, the selection process needs to be fully documented, with both the inclusion and exclusion criteria being specified in the protocol.

4.1. Inclusion and Exclusion criteria for Study Selection

Studies will be selected for the literature review if they meet the following inclusion criteria:
- The study looks at resource estimation within the domain of Web development. Studies can consider any facet of resource estimation, for example, effort estimation. Studies that use student data (as opposed to commercial data) will also be considered.
- The study describes the methodology, metrics, and datasets used for resource estimation.
- The study provides an empirical basis for its findings.

With regards to exclusion criteria, studies will be excluded if they:

- Do not focus on estimating a resource factor that is relevant to Web development.
- Do not provide an empirical basis to their findings.

The inclusion and exclusion criteria defined above are not set in stone and may be refined during the literature review process [7]. Any changes made to the inclusion and exclusion criteria naturally have to be documented.

4.2.   Initial Selection Process

With the inclusion and exclusion criteria decided upon, the primary studies identified by the search phase are screened. Their titles and abstracts are extracted and compiled into a list (for example in a Word document), and for those that are found relevant, a hardcopy is retrieved. In the situation that the title and the abstract are not sufficiently detailed to determine a paper's relevance, a hardcopy will be retrieved and used to make a decision. This selection process will be performed by both the student and their supervisor/s.

4.3.   Final Selection Process

In the final selection process, the hardcopies retrieved in the initial selection are analyzed in detail and if the study is still found to be relevant at this stage, it is added to the final reference library for the systematic literature review. Hardcopy analysis is also useful for any papers where there is a disagreement about inclusion or exclusion, between the students and their supervisor/s.

5. Study Quality Assessment

Assessing the quality of the studies in the final reference library is the next step in the systematic literature review process. This section will detail a quality assessment checklist that will provide a means to quantitatively assess the quality of the evidence presented by these studies. The conclusions drawn from a literature review are only as strong as the evidence they are based on, so compiling appropriate checklists to assess study "quality" is important [7]. As such, these checklists are not meant to be a form of criticism of any researchers' work.

Given the quantitative nature of software estimation, we expect to encounter only quantitative studies in our final reference library. Given that the items in this checklist are dependent on the primary studies selected for the systematic review [17] these items may be updated if the need arises. All changes to the quality assessment criteria will be documented.

Table 9 details the quality assessment checklist used to evaluate the primary studies. This checklist has been adapted from those compiled by Kitchenham [7], with each question utilizing the same answer scale, as summarized in Table 8.

Answer	Score
Yes	1
No	0
Partially	0.5

Table 8: Answer scale for questions in the quality checklists.

Using the above answer scale the higher the overall score a study obtains, the greater the degree with which this study addresses the research questions, and hence the greater its quality is.

No.	Question	Answer
1	Are the research aims clearly specified?	Yes/No/Partially
2	Was the study designed to achieve these aims?	Yes/No Partially
3	Are the prediction techniques used clearly described and their selection justified?	Yes/No/Partially
4	Are the variables considered by the study suitably measured?	Yes/No/Partially
5	Are the data collection methods adequately detailed?	Yes/No/Partially
6	Is the data collected adequately described?	Yes/No/Partially
7	Is the purpose of the data analysis clear?	Yes/No/Partially
8	Are the statistical techniques used to analyze the data adequately described and their use justified?	Yes/No/Partially
9	Were potential confounders suitably controlled for in the analysis?	Yes/No/Partially
10	Are the study findings credible?	Yes/No/Partially
11	Are negative results (if any) presented?	Yes/No/Partially
12	Do the researchers discuss any problems with the validity/reliability of their results?	Yes/No/Partially

Table 9: Quality assessment checklist for primary studies, adapted from [7].

6. Data Extraction

With the final set of primary studies decided upon and their quality assessed, the data extraction phase of the systemic literature review process can begin. The following section specifies the data extraction form created for recording data, and details the strategy for using them.

6.1. Data Extraction Forms

A data extraction form is designed to record all the relevant information from the primary studies required to answer the research questions. Table 11 represents the data extraction form used for this systematic literature review.

Data Item	Value	Supplementary Notes
Study Information Data		
Study ID	A unique identifier in the format: S<number>.	
Title		
Author(s)		
Year of publication		
Reference type	Journal/Conference/Thesis/Unpublished	
Publisher		
Data Relevant to Answering Research Questions		
Data characteristics	Industry(single/cross company)/Academia	
What methods/techniques were used for resource estimation?		
What factors are used as resource predictors?		
At what stage of the project were the predictors gathered?		
Are causal relationships between resource predictors accounted for?	Model/Technique: Yes/No	
If causal relationships are accounted for, are they probabilistic?	Model/Technique: Yes/No	
What metrics have been used to measure estimation accuracy?		
What accuracy did these methods/techniques achieve?		
Quality Assessment Checklist		

Are the research aims clearly specified?	Yes/No/Partially
Was the study designed to achieve these aims?	Yes/No/Partially
Are the prediction techniques used clearly described and their selection justified?	Yes/No/Partially
Are the variables considered by the study suitably measured?	Yes/No/Partially
Are the data collection methods adequately detailed?	Yes/No/Partially
Is the data collected adequately described?	Yes/No/Partially
Is the purpose of the data analysis clear?	Yes/No/Partially
Are the statistical techniques used to analyze the data adequately described and their use justified?	Yes/No/Partially
Were potential confounders suitably controlled for in the analysis?	Yes/No/Partially
Are the study findings credible?	Yes/No/Partially
Are negative results (if any) presented?	Yes/No/Partially
Do the researchers discuss any problems with the validity/reliability of their results?	Yes/No/Partially

Table 10: Data extraction form for quantitative studies.

6.2. Data Extraction Process

The above data extraction form will be used to extract data from the primary studies. The extraction process will be carried out by the student responsible for the literature review. Their supervisor will perform data extraction on a subset of these primary studies and their results compared to those of the student's for validation purposes. In a situation where data is difficult to understand or not clearly detailed in the study, the main author of the study will be contacted for clarification.

The data extracted from the primary studies will be saved as Microsoft Word 2007 documents with each study commanding its own file. The Word documents will be named according to the format *<main author name>_<year of publication>*. If more than one primary study has been published by the same author in the same year, then a unique identifier (e.g. a sequence number) will be added to the file name.

7. Data Synthesis

The data synthesis phase of the systematic literature review process involves compiling the data extracted from the primary studies so as to address each of the research questions. Data synthesized for each question will is tabulated facilitating any future analysis required.

7.1. Question 1

Questions 1, 1a, 1b, and 1c state:

What methods and techniques have been used for Web resource estimation?

What metrics have been used to measure estimation accuracy?

What (numerical) accuracy did these various methods/techniques achieve?

These questions deal with the methods/techniques used for Web resource estimation. All relevant data extracted from the primary studies is compiled in Table 13.

Study ID	Estimation Technique	Accuracy Measure Used	Estimation Accuracy Achieved	Notes (If Any)

Table 11: Summary of evidence on resource estimation methods/techniques.

7.2. Question 2

Questions 2, 2a, and 2b state:

What factors (e.g. effort, quality, size) have been investigated as resource predictors for Web development?

At what stage are these resource predictors gathered?

What are the most investigated resource predictors?

This set of questions looks at what Web development factors are used for resource estimation. Answering these questions will inform us as to which factors are considered most important for Web resource estimation, and when in the development cycle they are gathered. Table 14 summarizes resource predictor information, and Table 15 keeps track of how often each predictor encountered is utilized.

Study ID	Resource Predictors Used	Stage At Which Resource Predictors Are Gathered	Notes (If Any)

Table 12: Summary of web development factors used for resource estimation.

Resource Predictor	Frequency Used	Notes (If Any)

Table 13: Frequency with which each resource predictor has been utilized.

7.3. Question 3

Question 3 states:

What are the characteristics (single/cross-company, student/commercial projects) of the datasets used for Web resource estimation?

The final research question looks at the characteristics of the datasets used by the primary studies. Of key interest are whether single or cross-company data has been utilized and the setting in which the experiment was held. Information synthesized here may be of benefit when considering differences in estimation accuracy between different estimation techniques, and for that matter studies.

Study ID	Data Characteristics	Notes (If Any)

Table 14: Summary of the data characteristics of the primary studies.

8. Schedule for Review

The schedule for the systematic literature review process is summarized in the Gantt chart below.
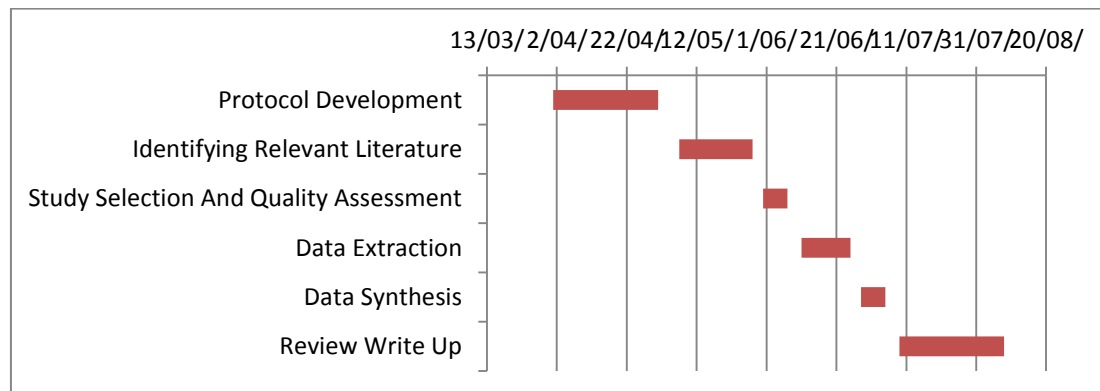

Figure 1: Schedule for systematic literature review.

9. References

1) N. Fenton, W. Marsh, M. Neil, P. Cates, S. Forey, and M. Tailor, *Making resource decisions for software projects,* Software Engineering, 2004. ICSE 2004. Proceedings. 26th International Conference, 23-28 May 2004 Page(s):397 – 406

2) B.W. Boehm, and R. Valerdi, *Achievements and Challenges in Cocomo-based software resource estimation,* Software, IEEE, Volume 25, Issue 5, Sept.-Oct. 2008 Page(s):74 – 83

3) E. Mendes, *A Comparison of Techniques for Web Effort Estimation*, Empirical Software Engineering and Measurement, 2007. ESEM 2007. First International Symposium, 20-21 Sept. 2007 Page(s):334 – 343

4) Agena: Bayesian Network and Simulation Software for Risk Analysis and Decision Support, http://www.agenarisk.com/

5) C. Kiran, The Web Development Industry Is Expected To Grow Over 20 By 2010, http://www.articler.com/23205/The-Web-Development-Industry-Is-Expected-To-Grow-Over-20-By-2010.html

6) P. Umbers, and G. Miles, *Resource estimation for Web applications,* Software Metrics, 2004. Proceedings.10th International Symposium, 14-16 Sept. 2004 Page(s):370 – 381

7) B. Kitchenham, *Guidelines for performing systematic literature reviews in software engineering (version 2.3)*, Software Engineering Group, School of Computer Science and Mathematics, Keele University and Department of Computer Science, University of Durham, July 2007.

8) M. Petticrew, and H. Roberts, *Systematic Reviews in the Social Sciences: A Practical Guide*, Blackwell Publishing, 2006.

9) K. Agusa, Software Engineering Evolution, 7th International Workshop on Principles of Software Evolution, 2004, Proceedings, Page(s): 3 – 8.

10) L. Baresi, and S. Morasca, *Three empirical studies on estimating the design effort of Web applications*, ACM Transactions on Software Engineering and Methodology (TOSEM) archive, Volume 16, Issue 4, September 2007.

11) E. Mendes, and N. Mosley, *Comparing effort prediction models for Web design and authoring using boxplots*, .24th Australasian Computer Science Conference, ACSC 2001, Proceedings, 29 Jan.-04 Feb. 2001, Page(s): 125 – 133.

12) E. Mendes, *Investigating metrics for a development effort prediction model of Web applications*, Australian Software Engineering Conference, 2000, Proceedings, 28-29 Apr. 2000, Page(s): 31 – 41.

13) E. Mendes, N. Mosley, and S. Counsell, *Using an engineering approach to understanding and predicting Web authoring and design*, 34th Annual Hawaii International Conference on System Sciences, 2001, Proceedings, 3-6 Jan. 2001.

14) S. Tilley, Preliminary results from a case study of effort estimation for net-centric applications at the undergraduate level, 14th Conference on Software Engineering Education and Training, 2001, Proceedings, 19-21 Feb. 2001, Page(s): 56 – 63.

15) M. Morisio, I. Stamelos, V. Spahos, and D. Romano, *Measuring functionality and productivity in Web-based applications: a case study*, Sixth International Software Metrics Symposium, 1999, Proceedings, 04-06 Nov. 1999, Page(s): 111 – 118.

16) M. Jorgensen, and M. Shepperd, *A Systematic Review of Software Development Cost Estimation Studies*, IEEE Transactions on Software Engineering, Volume 33, Issue 1, Jan. 2007, Page(s): 33-53.

17) M. Riaz, M. Sulayman, N. Salleh, and E. Mendes, *Experiences Conducting Systematic Reviews from Novices' Perspective*, 14th International Conference on Evaluation and Assessment in Software Engineering, EASE 2010, 12-13 Apr. 2010.