

# Chapter 6

## A Systematic Review of Attributes and Techniques for Open Source Software Evolution Analysis

**Munish Saini**

 <https://orcid.org/0000-0003-4129-2591>

*Guru Nanak Dev University, India*

**Kuljit Kaur Chahal**

*Guru Nanak Dev University, India*

### ABSTRACT

*Many studies have been conducted to understand the evolution process of Open Source Software (OSS). The researchers have used various techniques for understanding the OSS evolution process from different perspectives. This chapter reports a meta-data analysis of the systematic literature review on the topic in order to understand its current state and to identify opportunities for the future. This research identified 190 studies, selected against a set of questions, for discussion. It categorizes the research studies into nine categories. Based on the results obtained from the systematic review, there is evidence of a shift in the metrics and methods for OSS evolution analysis over the period of time. The results suggest that there is a lack of a uniform approach to analyzing and interpreting the results. There is need of more empirical work using a standard set of techniques and attributes to verify the phenomenon governing the OSS projects. This will help to advance the field and establish a theory of software evolution.*

### 1. INTRODUCTION

Due to the rising dominance of Open Source Software (OSS) in the software industry; not only are practitioners, but researchers as well as academicians also keen to understand the OSS development and evolution process. OSS development involves various stakeholders ranging from contributing volunteers

DOI: 10.4018/978-1-6684-3702-5.ch006

to commercial software vendors. There is need to understand the OSS development model in general and OSS evolution in particular so that the evolution process can be improved, if need be, for the future systems.

OSS evolution has attracted a lot of attention in the last decade. Easy and free availability of data on open source projects has resulted in a splurge of studies in this domain. As a result, the number of empirical studies related to OSS is much more in number in comparison to other topics in the field (Stol and Babar, 2009). Various methods have been employed in the past for analysis and prediction of OSS evolution. It is necessary to systematically summarise the empirical evidence obtained on these methods from the existing literature so that it is easy to comprehend the research work in this area, and reveal gaps in the existing work. As per the existing work in this direction, a few studies focusing on the survey of literature in the domain have been published. Fernandez- Ramil et al. (2008) discuss, in an informal way, a small sample (seven in numbers) of OSS evolution studies. Breivold et al. (2010) carry out a systematic literature review of OSS evolution studies (41 in numbers) focusing only on the evolvability characteristic of OSS systems. Syeed et al. (2013) follows a systematic literature review protocol to analyze studies on OSS Evolution. They present review of 101 research papers but their focus is on a limited set of categories of studies. Stol and Babar (2009) reviewed empirical studies reported in four International OSS conferences to assess quality of the papers from the perspective of the way they report the empirical research in OSS. Unlike the present study, their target is not review of studies on OSS evolution but assessment of quality of empirical research papers involving OSS systems. This chapter presents a systematic literature review of an extensive list of research papers published on the subject between the period of 1997 and 2016.

A number of research publications on OSS evolution have explored the phenomenon from different dimensions using different approaches. Broadly two dimensions are taken: Evolution in OSS structure, and Evolution in OSS community. Software structure exploration includes source code analysis, version history analysis, and repository information analysis. Community structure exploration includes social network analysis. Both the dimensions cannot be isolated from each other. They are useful when put together, and complement each other in answering questions regarding the OSS development and evolution process. Analyzing the links between the software structure and the developer community helps in improving software evaluation and quality.

In this chapter, we report a meta-data analysis on comprehensive review on OSS evolution published in the time period of 1997 to 2016 along with discussion on the project attributes and techniques used for analyzing software evolution (Chahal and Saini, 2016a; 2016b).

The rest of the paper is organized as follows: Section 2 presents the research questions that are addressed in this systematic review and the research criteria followed in this study for selection of primary studies. Section 3 presents the answers to the research questions identified in this work. Section 4 gives conclusions and future directions obtained from this systematic review.

## **2. RESEARCH METHODOLOGY**

The review process follows a systematic review protocol (Kitchenham, 2007) so as to reduce the research bias. The review process included the following steps: 1) Defining the research questions, 2) Choosing a search strategy and study selection criteria, and 3) Data Extraction and Synthesis.

21 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/a-systematic-review-of-attributes-and-techniques-for-open-source-software-evolution-analysis/294460](http://www.igi-global.com/chapter/a-systematic-review-of-attributes-and-techniques-for-open-source-software-evolution-analysis/294460)

## Related Content

---

### Performance Analysis of a Web Server

Jijun Lu and Swapna S. Gokhale (2009). *Software Applications: Concepts, Methodologies, Tools, and Applications* (pp. 366-379).

[www.irma-international.org/chapter/performance-analysis-web-server/29397](http://www.irma-international.org/chapter/performance-analysis-web-server/29397)

### A Model Based on Data Envelopment Analysis for the Measurement of Productivity in the Software Factory

Pedro Castañeda and David Mauricio (2022). *Research Anthology on Agile Software, Software Development, and Testing* (pp. 792-818).

[www.irma-international.org/chapter/a-model-based-on-data-envelopment-analysis-for-the-measurement-of-productivity-in-the-software-factory/294495](http://www.irma-international.org/chapter/a-model-based-on-data-envelopment-analysis-for-the-measurement-of-productivity-in-the-software-factory/294495)

### Frequency-Reponse Masking Techniques

Ljiljana Milic (2009). *Multirate Filtering for Digital Signal Processing: MATLAB Applications* (pp. 295-315).

[www.irma-international.org/chapter/frequency-reponse-masking-techniques/27219](http://www.irma-international.org/chapter/frequency-reponse-masking-techniques/27219)

### Engineering Financial Enterprise Content Management Services : Integration and Control

Dickson K. W. Chiu, Patrick C. K. Hung and Kevin Kwok (2010). *International Journal of Systems and Service-Oriented Engineering* (pp. 86-113).

[www.irma-international.org/article/engineering-financial-enterprise-content-management/44688](http://www.irma-international.org/article/engineering-financial-enterprise-content-management/44688)

### Enforcing Modeling Guidelines in an ORDBMS-based UML-Repository

N. Ritter and H.P. Steiert (2002). *Optimal Information Modeling Techniques* (pp. 150-162).

[www.irma-international.org/chapter/enforcing-modeling-guidelines-ordbms-based/27833](http://www.irma-international.org/chapter/enforcing-modeling-guidelines-ordbms-based/27833)