

On the Efficacy of Keyword Searches to Find Meaningful Architectural Knowledge in Open-Source Software Mailing Lists

Bachelor Thesis for Computing Science

Andrew Lalis

Supervised by Mohamed Soliman

May 10, 2022



**university of
 groningen**

**faculty of science
and engineering**

Abstract

This section will contain a brief overview of the research and conclusions.



Contents

Abstract	1
Introduction	3
Background	3
Other	3
Background Information	4
Research Questions	5
Methodology	6
Results	7
Conclusion	8



Introduction

In this paper, we'll explore the efficacy of using targeted keyword search queries to find architectural knowledge in mailing lists for open-source software projects. More simply put, we'll build tools, collect data, and analyze that data to qualitatively determine how effective certain keyword-based search queries are at finding useful information in large sets of emails.

Background

Test

Other

test!



Background Information

This section contains some background information about what research already exists on this topic, and what knowledge it provides.



Research Questions

The main research question that this paper attempts to answer is summarized in the following question:

What is the effectiveness of using keyword search queries to find architectural knowledge in open-source software mailing lists?

In addition to the main question we're attempting to answer, this paper will also discuss several other possible questions and answers that may be obtained using the data originally gathered for the main question.

1. Does there exist a relationship in the order in which architectural decisions are discussed, chronologically in an email thread? Is there significance in the order in which architectural decisions are made?
2. Is there a relationship between the content of discussion in emails, and related issues in issue/ticket boards such as [JIRA](#) and [GitHub issues](#)?
3. How can we use data gathered in this research to improve our search queries?



Methodology

This section will explain the process that was followed to obtain data.



Results

This section will show visualizations and aggregate data for results.



Conclusion

This section will answer the research questions and discuss further research.

