

**SENTIRESEARCH: LEXICON-BASED WEB APPLICATION FOR
INDONESIAN SENTIMENT ANALYSIS**

by

Sabika Amalina (黄秀英)

A Thesis

*Submitted to the Faculty of Nanjing Xiaozhuang University
In Partial Fulfillment of the Requirements for the degree of*

Bachelor of Software Engineering



School of Information Engineering

Fangshan, Nanjing

May 2020

DECLARATION OF ORIGINALITY

I, the undersigned below:

Name : Sabika Amalina

Student ID : 16523002

Hereby declared that the thesis I wrote with the title: SENTIRESEARCH:
LEXICON-BASED WEB APPLICATION FOR INDONESIAN SENTIMENT
ANALYSIS.

1. Is truly a research written and conducted purely by myself, not copying from other published researches, and also not a result of plagiarism.
2. I will allow Nanjing Xiaozhuang University and Universitas Islam Indonesia to manage and keep the copy of this thesis, to be used as they deem necessary.

I made this statement of declaration with fully responsibility, and I'm willing to accept any consequence according to the rules and regulations should the statement above proved to be wrong in any way.

Nanjing, May, 2020

Sabika Amalina

To: Dean Xiangjun Zhao
School of Information Engineering

This thesis, written by Sabika Amalina (黄秀英), and entitled SENTIRESEARCH: LEXICON-BASED WEB APPLICATION FOR INDONESIAN SENTIMENT ANALYSIS, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this thesis and recommend that it be approved.

Qing Li

Wan Li Song

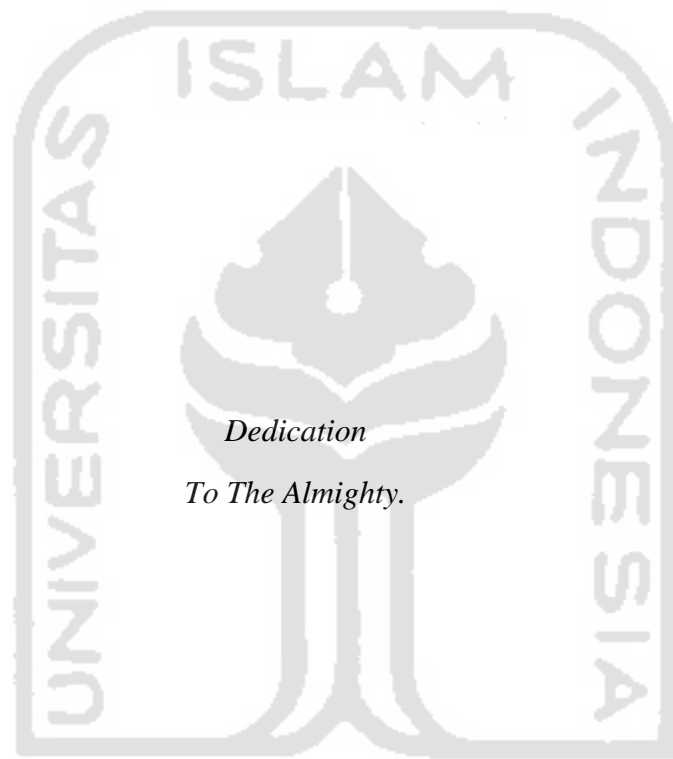
Hai Yong Wu

Date of Defense: May 19, 2020

The thesis of Sabika Amalina is approved.

Dean Xiangjun Zhao
School of Information Engineering

Zhu Ming Hui
Chief of Foreign Affairs Office



*Dedication
To The Almighty.*

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

ACKNOWLEDGMENTS

In the name of Allah The Most Gracious and The Most Merciful. First and foremost, all praises and thanks to Allah, for His mercy and guidance. It is only because of His endless blessings that enable me to complete this research work as my bachelor's thesis despite of my flaws, limitations, and uncertainties we might face due to COVID-19 pandemic. Peace and prayers be upon His Final Prophet and Messenger, Muhammad, the ideal role model of human beings who has guided us from the darkness to the lightness, and from stupidity to cleverness.

I would like to express my deep and sincere gratitude to my beloved family members: my loving parents, Drs. Turmuzi and Siti Sayidah Khumairoh; my supportive siblings, Mas Hafidz, Mba Atika, and twin sister Sabila; and my whole big family in Bantul, Yogyakarta. Thank you for the generous support and love they have provided me throughout my entire life, particularly through my journey of pursuing my bachelor's degree in Universitas Islam Indonesia and Nanjing Xiaozhuang University.

In the writing of this thesis, I would like to express immeasurable appreciation for the help and support to the following persons who in one way or another have contributed in making this thesis possible:

1. Dr. Haiyong Wu, as my thesis advisor, who has guided me from the beginning until the completion of the thesis.
2. Prof. Fathul Wahid, S.T., M.Sc., Ph.D., as the Rector of Universitas Islam Indonesia; Hendrik, S.T., M.Eng., as the Head of Department of Informatics of Universitas Islam Indonesia; Taufiq Hidayat S.T., M.C.S., as my second thesis advisor; and respectful lecturers in Department of Informatics, Faculty of Industrial Engineering, Universitas Islam Indonesia.
3. Wangjing and Sophie Mou 老师, who always take care of the international students in Nanjing Xiaozhuang University, especially the students of Software Engineering.

4. The best people I have ever met in my university life: Puspita Dewi, Ulfa Amalia, Assyifa Narulita, Fionna Saphira, Sari Kurnia, and Rizal Hamdan. No words could express how fortunate I am to have you as my support system.
5. My comrade-in-arms, the students of Software Engineering double degree program class of 2020 and Hexadecima. We are in this together.
6. The people in all of the companies and organizations I have ever taken part in: Kalikesia, the growing startup company; Gojek, the pride of Indonesia; Marketing and Communication of FTI UII; Indonesian Student Association in Nanjing; and many other more.
7. My kindhearted seniors, Oktri Nanda, Hanifa, and Inggih Wicaksono, who have helped me in finding solution and solving problem for to my thesis work.
8. The shining stars and shining diamonds, whose smile could light up my days and whose dream always make me motivated to be the better version of myself day by day.
9. Last but not least, to myself who have done her best and those who cannot be mentioned one by one, who have helped and support me to finish this thesis.

May God bless all of those mentioned above for their sacrifices to become merciful deeds. Along with the completion of this thesis, the author hopes that this small work could be beneficial for the others and could be developed for the better in the future.

Yogyakarta/Nanjing, 18 May 2020

Sabika Amalina



TABLE OF CONTENTS

TABLE OF CONTENTS.....	i
LIST OF TABLES.....	iv
LIST OF FIGURES	v
ABSTRACT.....	1
CHAPTER 1. INTRODUCTION.....	2
1.1 The Background.....	2
1.2 Problem Identification.....	3
1.3 Purpose of the Study	4
1.4 Limitation of the Study	4
1.5 The benefit of the Study.....	4
1.6 Research Methodology.....	5
1.7 Writing Structure.....	5
CHAPTER 2. REVIEW OF LITERATURE.....	7
2.1 Fundamental Theory	7
2.1.1 Qualitative Research.....	7
2.1.2 Natural Language Processing	8
2.1.3 Sentiment Analysis.....	9
2.1.4 Text Preprocessing.....	12
2.1.5 Python	13
2.1.6 Django Framework	13
2.1.7 Prototype.....	14
2.2 Related Works	14
CHAPTER 3. ANALYSIS AND DESIGN	16
3.1 Research Framework.....	16

3.1.1	Problem Identification	16
3.1.2	Literature Study	17
3.1.3	Data Collection	17
3.1.4	Software Development Life Cycle.....	18
3.2	System Requirement Analysis	19
3.2.1	Functional Requirement.....	19
3.2.2	Non-Functional Requirement.....	22
3.2.3	Actor Identification.....	Error! Bookmark not defined.
3.3	System Design.....	22
3.3.1	Use Case Diagram.....	23
3.3.2	Site Map	23
3.3.3	Database Schema Design.....	27
3.3.4	Architectural Design	27
3.3.5	System Low Fidelity Prototype	28
3.4	Sentiment Analysis Methodology	33
3.4.1	Text Preprocessing	34
3.4.2	InSet Lexicon Dictionary-Based.....	35
CHAPTER 4. IMPLEMENTATION AND TESTING		36
4.1	System Development Specification	36
4.2	Back-End Implementation.....	37
4.2.1	Installed Modules and Packages	37
4.2.2	URL Configurations.....	38
4.3	Front-End Implementation	41
4.3.1	Landing App User Interface.....	42
4.3.2	Repository App User Interface.....	45
4.3.3	Sentiment App Dashboard User Interface.....	52
4.4	Sentiment Analysis Implementation	52
4.4.1	Text Preprocessing	52
4.4.2	Sentiment Analysis.....	56
4.5	Testing.....	59

4.5.1	Landing App Testing.....	59
4.5.2	Repository App Testing.....	60
4.5.3	Sentiment App Testing.....	62
CHAPTER 5. CONCLUSION AND RECOMMENDATION		63
5.1	Conclusion.....	63
5.2	Recommendation.....	64
APPENDIX A. SURVEYS.....		65
LIST OF REFERENCES.....		72



LIST OF TABLES

Table 3.1 Actor Identification.....	19
Table 3.2 Functional Requirement.....	21
Table 3.3 Non-Functional Requirement	22
Table 4.1 System Development Specification	36
Table 4.2 Modules and Packages.....	37
Table 4.3 Landing App URLconf.....	39
Table 4.4 Repository App URLconf.....	39
Table 4.5 Sentiment App URLconf	41
Table 4.6 Example of Sentence before Preprocessing.....	55
Table 4.7 Preprocessed Sentence.....	56
Table 4.8 Assign Polarity Score Result	57
Table 4.9 Sentiment Score	57
Table 4.10 Landing App Testing Result.....	59
Table 4.11 Repository App Testing Result.....	60
Table 4.12 Sentiment App Testing Result	62

LIST OF FIGURES

Figure 2.1 Sentiment Analysis Approaches.....	11
Figure 3.1 Research Framework.....	16
Figure 3.2 Prototype Method – Software Development Life Cycle.....	18
Figure 3.3 Use Case Diagram.....	23
Figure 3.4 Admin Sitemap.....	24
Figure 3.5 Project Manager Sitemap.....	25
Figure 3.6 Researcher Sitemap.....	26
Figure 3.7 General User Sitemap.....	26
Figure 3.8 Database Schema.....	27
Figure 3.9 Architecture Design.....	28
Figure 3.10 Landing Page Mockup.....	29
Figure 3.11 Authentication Mockup.....	30
Figure 3.12 Register and Join Workspace Mockup.....	30
Figure 3.13 Workspace Dashboard Mockup.....	31
Figure 3.14 Research Details Mockup.....	31
Figure 3.15 Research Table Mockup.....	32
Figure 3.16 Research Form Mockup.....	32
Figure 3.17 Sentiment Analysis Insight Chart Mockup.....	33
Figure 3.18 Sentiment Analysis Flowchart.....	34
Figure 3.19 Text Processing Flowchart.....	34
Figure 3.20 InSet Lexicon Dictionary in Excel.....	36
Figure 4.1 Django Project Structure.....	42
Figure 4.2 Landing Page User Interface.....	43
Figure 4.3 Sign in Page User Interface.....	44
Figure 4.4 Sign up Page User Interface.....	44

Figure 4.5 Workspace Register Page User Interface	45
Figure 4.6 Workspace Join Request Page User Interface	45
Figure 4.7 Workspace Dashboard Page	46
Figure 4.8 Research Project Management Page User Interface.....	46
Figure 4.9 Research Project Details Page User Interface	47
Figure 4.10 Project’s Respondent Management Section User Interface	47
Figure 4.11 Project’s Question Management Section User Interface.....	48
Figure 4.12 Project’s Question and Answer Page User Interface.....	49
Figure 4.13 Research Respondent Management Page User Interface	49
Figure 4.14 Research Respondent Details Page User Interface.....	50
Figure 4.15 Workspace Details Page User Interface	51
Figure 4.16 Workspace Join Request Page User Interface	51
Figure 4.17 Sentiment App Dashboard User Interface	52
Figure 4.18 Modules and Packages Imported.....	53
Figure 4.19 File and Library Prepared for Preprocessing.....	54
Figure 4.20 Function of Text Preprocessing.....	55
Figure 4.21 Function of Sentiment Scoring.....	57
Figure 4.22 Function of Sentiment Classification	58
Figure A.1 Type of research Researcher mostly do.....	65
Figure A.2 Research method that have ever been done.....	66
Figure A.3 Role that would likely to request to do research	66
Figure A.4 Frequency of Qualitative Research.....	66
Figure A.5 Researcher needed in one research project.....	67
Figure A.6 Availability of research repository	67
Figure A.7 Type of platform to design Qualitative Research.....	68
Figure A.8 Ways to take note or record an interview	68
Figure A.9 Challenge in research synthesis	68

Figure A.10 The importance of tools.....	69
Figure A.11 The importance of features.....	69
Figure A.12 Respondents' Gender.....	70
Figure A.13 Respondents' Age.....	70
Figure A.14 Working period as a Researcher.....	70
Figure A.15 Number of employees in the company.....	71
Figure A.16 Number of Researchers in the company.....	71



ABSTRACT

Today we are living in the “age of customer”, where customer or user knowledge and perception is a key for running successful business. Thus, many companies in the world have the role of researcher to dive deep and do research specifically for the customers or users. Analyzing qualitative data is one of the main tasks that every researcher has to tackle and it can even be laborious for advanced market research firms. Therefore, the researchers find the research results are difficult to analyze because the responses can be very haphazard understanding and there might be bias over feedbacks. To overcome this issue, Sentiment Analysis as part of Natural Language Processing is applied in the development of web-based applications that allow for visualization of current sentiments associated with a keyword of open question results from in-depth interview narrative. This research used lexicon dictionary-based sentiment analysis with research respondents’ answers or feedbacks as the dataset. The previous research also has proven to have better Sentiment Analysis result with using InSet Lexicon with the range of -5 to +5. The result of this research showed that this system can benefit the research team with a complete end-to-end system for Sentiment Analysis of customer feedbacks in Bahasa Indonesia from interview session along with the data visualization.

Keywords: Sentiment Analysis, Natural Language Processing, Qualitative Research, Interview.

CHAPTER 1.INTRODUCTION

1.1 The Background

Sentiment Analysis is a sub-field of Natural Language Processing (NLP) that tries to identify and extract opinions within a text, can be called with Opinion Mining. The main purpose of sentiment analysis is to determine the perspective, sentiments, evaluations, attitudes, and emotions of certain speakers/writers based on the computational treatment of subjectivity in a text. Sentiment Analysis is also useful for practitioners and researchers, especially in fields like marketing, psychology, economics, and politics, which rely a lot on human-computer interaction data information.

Revealing and getting the insight from sentiment analysis in user behaviors, needs, and motivations to design products and services that provide value is the crux of user experience research. Once the analysis is performed correctly, sentiment analysis has a huge impact on business. As such, the role of the user experience (UX) researcher is becoming more specialized and more in-demand. A 2017 report into the user research industry found that 81% of executives surveyed agreed that user research makes their company more efficient. Also, 86% of executives believe that user research improves the quality of their products and services (Moebius, 2017).

User experience research is the systematic investigation of users and their requirements in order to add context and insight into the process of designing the user experience. UX research employs a variety of techniques, tools, and methodologies to reach conclusions, determine facts, and uncover problems, thereby revealing valuable information that can be fed into the design process.

One of the methodologies essentially used by UX researchers is an in-depth interview. A user or in-depth interview is a UX research method during which a researcher asks one user questions about a topic of interest (e.g., use of a system, behaviors and habits) with the goal of learning about that topic. Unlike focus groups, which involve multiple users at the same time,

user interviews are one-on-one sessions (although occasionally several facilitators may take turns asking questions).

UX Interviews tend to be a quick and easy way to collect user data, so they are often used, especially in Lean and Agile environments. They are closely related to journalistic interviews and to the somewhat narrower and more formal HCI method called the critical incident technique, which was introduced in 1954 by John Flanagan.

Humans are subjective creatures, and opinions are important. Therefore, the author wants to implement sentiment analysis to help researchers collecting insights within the research outcomes like notes or scripting for an in-depth interview, questionnaires, and diary study methodology. Thus, this study aims to develop a web-based application system that implements specifically sentiment analysis technology. The application system is expected to be able to help UX Researcher as well as Project Manager and other employees in the qualitative research projects.

Based on the explained background above, the author intended to research by taking the thesis paper entitled: “Web-based Sentiment Analysis Application System for Qualitative Research Project Management”.

1.2 Problem Identification

Based on the background that has been explained above, the author identified some problems as motives for the development of this application. The problems are as follow:

1. How to design the web-based application system that could store and manage qualitative research narratives, then as well as to analyze the sentiment from research results?
2. How to develop the web-based application system specified for sentiment analysis?
3. How does this web-based application system benefit the research team in a company to help them schedule and manage their research projects?

1.3 Purpose of the Study

The purpose of this application system development are listed as follow:

1. This study aims to develop a web-based application system for the research team to get sentiment analysis as a result from the qualitative research notes.
2. This web-based application is designed to improve collaboration within research team.
3. This web-based application is intended to be part of the user experience research system in the information technology industry.

1.4 Limitation of the Study

Some limitations of this application system are listed as follow:

1. This application applied to the web-based platform.
2. The system only focuses on qualitative research project specified in in-depth interview activity.
3. This application is intended for the research team member that consists of Researchers, Project Manager, or other employees who work in Bahasa Indonesia.
4. The users are limited to those who work in a certain company that has been registered to the system.

1.5 The benefit of the Study

The benefits of this application system are listed as follow:

1. This application can help users store qualitative research notes.
2. This application can show the dashboard of sentiment analysis data results.
3. This application can manage and schedule research projects.
4. This application can generate word count from the qualitative research notes.

1.6 Research Methodology

The software development process is a very complicated task if it is done without any proper step by step procedure. Then, in order to make the software development processes simple and systematic, the Software Development Life Cycle (SDLC) came into existence. The SDLC defines the framework that includes various activities and tasks to be carried out throughout the software development process. Various SDLC methods are used in the software development process, heaving their advantages and disadvantages.

In this research, the author implemented SDLC Prototype method to the development process. This method emphasizes on the iterative design process that enables the target user to try the pre-release or sample of the system before the development process occurs. During the process, the prototype is built, tested, and then reworked based on user's feedbacks and suggestions as necessary until an acceptable outcome is achieved and finally proceed to the development phase.

1.7 Writing Structure

The systematic writing is made to facilitate the writing of the final project report. Thus, the systematic writing of this thesis is divided into six chapters, with the explanation for each chapter, as follows:

CHAPTER I INTRODUCTION

This chapter discusses the background of the study, problem identification, purposes, limitations, benefits of the study, research methodology, and also schedules of this study.

CHAPTER II THEORETICAL BASIS

In this chapter discusses the basic theories that support and relate to this research.

CHAPTER III ANALYSIS AND DESIGN

This chapter discusses the analysis and design of the system. Analysis of the current system, analysis of problem, implementation of the methodology, problem solving with analyze the system needs and conducting the system design to generate a new system.

CHAPTER IV IMPLEMENTATION AND TESTING

This chapter discusses the result of the analysis and design into code to build a working application, and testing to find out whether the application is working or still needs improvements.

CHAPTER V CONCLUSION AND RECOMMENDATION

This chapter discuss the conclusion of this thesis and suggestions for further research.



CHAPTER 2. REVIEW OF LITERATURE

2.1 Fundamental Theory

2.1.1 Qualitative Research

Qualitative research is a research method that tends to focus on obtaining data from groups of people or individuals through an open-ended and conversational communication. This type of research attempts to get a deep understanding of how things came to be the way they are in our social world. The qualitative research data focuses on feedbacks, opinions, and experiences data that cannot be adequately expressed numerically.

Many researchers working in the social sciences like psychology, sociology, anthropology or anyone who are interested in studying human behavior often find it difficult to explain and describe human behavior in quantifiable, measurable terms. Therefore, a qualitative research process will follow the data collection methodological approach that a research team decides to adopt. Different approaches also involve different sets of assumptions about what sorts of information (or knowledge) are important (Hancock, 2006).

One of the mostly used qualitative data collection methods is In-Depth Interview (IDI) or simply an interview. IDI method is considered as a discovery-oriented method that aims to deeply explore the research participant or respondent's point of view, feelings, and perspectives. This method is closely related to the user experience approach.

UX(UX) is adopted by the human-computer interaction (HCI) community, which in practice, encompasses various aspects from the end-user interaction with the company, its services, and its products (Norman & Nielsen, 2016). The UX itself is a unique combination of various elements, such as the product and internal states of the user (e.g., mood, expectations, active goals), which extend over time with a definitive beginning and end. The outcome of this process is the experience.

2.1.2 Natural Language Processing

Natural language processing(NLP) is one branch of application of artificial intelligence in which computer is expected to be able to process text and provides results as desired (Indra & Hartati, 2014). In principle, natural language is represented in the form of sound or speech, as well as text or writing. With the applications of NLP, computers are expected to understand human language. According to certain research (Indra & Hartati, 2014), some of the implementations of NLP in our lives include:

1. Spelling and Grammar Correction

The application of spelling and grammar correction has been commonly found as auto-correction on smartphones.

2. Word Prediction

One application of word prediction often encountered is in the search engine when typing certain words in the search engine, and then automatically, the search engine will display the prediction of the next word that we have typed.

3. Text Summarization

Text summarization is one of the applications of NLP that enables the computer to provide a summary of a text document.

4. Information Retrieval

Information retrieval is one of the applications of NLP that we usually encounter in search that makes it easier for the search engine to find documents that match exactly as the query given by the user.

5. Sentiment Analysis

This sentiment analysis is one of NLP's abilities that can detect a positive, negative, or neutral comment.

6. Text Categorization

This text categorization will categorize a group of documents into one or several predefined categories.

7. Topic Modeling

Topic modeling is one method used to look for patterns or relationships between words in a document. This pattern and relationship will later be used to determine the topic of a particular document.

8. Question Answering

Question answering is one of the applications of NLP, and it is currently being developed. Chatbot application is often encountered nowadays as one of the implementations of the NLP. Question answering enables the application system to provide responses or answers to questions submitted by users.

9. Information Extraction

Information extraction is the application of NLP, which extracts essential words in a document and puts them in table data. In the application of information extraction, it also requires the name entity recognition.

10. Machine Translation

Machine translation is one of the applications of NLP that is quite often used in daily life, especially in the field of education. For example, google translate, which uses the implementation of this machine translation.

2.1.3 Sentiment Analysis

Sentiment Analysis (SA) is a sub-field of NLP that tries to identify and extract opinions within a text, can be called with Opinion Mining (Liu et al., 2005). The main purpose of sentiment analysis is to determine the perspective, sentiments, evaluations, attitudes, and emotions of particular speakers/writers based on the computational treatment of subjectivity in a text. Sentiment Analysis is useful for professionals and researchers, especially in fields like marketing, psychology, economics, and politics, which rely a lot on human-computer interaction data information.

SA is one of the hottest topics currently studied in NLP, and it has gained much popularity in

recent times due to the importance of the insights from user feedback that is crucial to improve customers' satisfaction. SA deals with the mood or opinion detection within a text of target topic or event using NLP, text analysis, and computational linguistics. Many research works (Andreevskaia et al., 2007; Jurek et al., 2015) have been done on SA, which focuses on classifying the text into a negative or positive category as compared to the emotion.

Knowledge-based approach is used to classify the sentence into types of emotions, which are happiness, joy, anger, surprise, sadness, fear, and anticipation, and non-emotion or neutral (Aman & Szpakowicz, 2007). A research has identified confusion, vigor, anger, fatigue, depression, and tension in the text other than these six emotions (Bollen et al., 2011). Also, Lexicon based approaches are found to perform better than machine learning approaches to classify the tweets data that can be extracted using a web crawler (Asghar et al., 2014; Westling et al., 2014).

Sentiment analysis can be applied to the voice of the customer materials in business needs. The major application of sentiment analysis can be found into three categories:

- a. Brand Reputation Monitoring: Sentiment Analysis is used by a specific company to see how its brand or product has been received by the consumer, user, or public in general. The insights gathered from the public's feedbacks are essential to better branding.
- b. Customer Service and Experience: Customer service agents classify incoming mail into 'urgent' and 'non-urgent' to be able to deliver a quick response and act in dealing with the customer.
- c. Market Research and Analysis: Opinion mining plays a crucial role in business intelligence that is in increasing demand for the company. The research, including benchmarking brand competitors and many more (Aroomogan, 2015).

There are three approaches for categorizing sentiment analysis which are: (a) Machine Learning based approach, (b) Lexicon based approach and (c) Hybrid approach. The approaches can be seen from **Error! Reference source not found.**(Sadia et al., 2018).

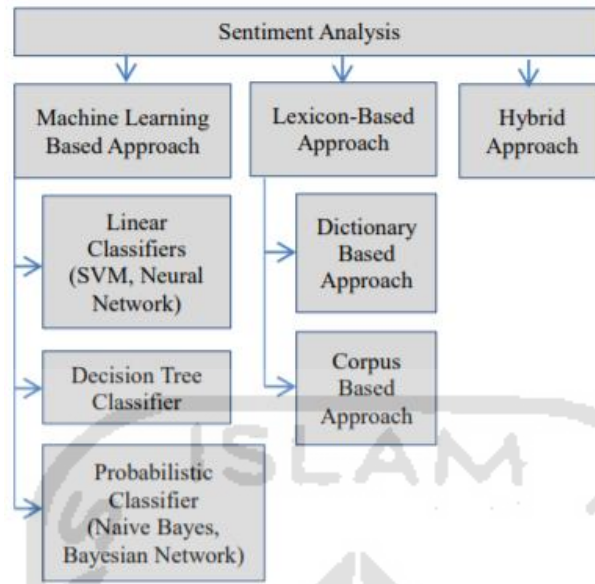


Figure 2.1 Sentiment Analysis Approaches

1. Machine Learning-Based Approach

Machine Learning is categorized as a supervised learning method which use the algorithms to train the classifier from manually labeled data. This approach classifies the text as positive, negative or neutral using Machine Learning classification algorithms and linguistic features.

2. Lexicon-Based Approach

This approach utilizes a sentiment lexicon to describe the polarity; whether it is positive, negative and neutral, of a textual content. This approach is considered more understandable and can be implemented easier compared to machine learning-based approach. This approach is further classified as:

- a. Dictionary-based method: It uses an online dictionary as well as a small set of seed words. The strategy here is a set of words with their known orientations that are collected, and then online dictionaries are searched to find the possible or nearest synonyms and antonyms. Later on, the sample will be classified based on the presence of signaling sentiment words. The Dictionary-based method is applied in this research of developing a web-application sentiment analysis with Lexicon-Based Approach.

b. Corpus-based method: This method uses corpus data to classify sentiment words. Even though the corpus-based not as effective as a dictionary-based method, it is a convenience to find the domain, context, and sentiment of specific of the words against the corpus data. The algorithm in this method will have access to sentiment labels and the context of the word or sentence.

3. Hybrid Approach

Hybrid is a combination of both Machine Learning and lexicon-based approaches. This combination of approaches has done by many researchers.

2.1.4 Text Preprocessing

In the field of Text Mining, data preprocessing used for extracting non-trivial and knowledge from unstructured data in the form of text. Information Retrieval (IR) is mostly a matter of deciding which information or documents should be retrieved that match the most as a user's need for information. The information that user needs is represented by a query that contains one or more search terms, also with some additional information such as the weight of each word. The decision of which data or information should be retrieved is made by comparing the terms of the query with the important words or phrases that appears in the text or document. The decision could be binary (retrieve/reject), or it could involve estimating the degree of relevance that the document has to query.

Before the selected data or information retrieved from the documents, data preprocessing has to be done due to the structural variants within the word that appear in the documents or the queries. The data preprocessing techniques are applied to the set of target data to reduce the data size, and that will increase the effectiveness of the IR system. The phases of Text Preprocessing include Tokenization, Stemming, and Stopword Removal.

2.1.5 Python

Python is known to be a high-level programming language that is widely used in the current software engineering industry created by Guido van Rossum. The design philosophy of Python emphasizes more on code readability, and the syntax allows programmers to have fewer lines of code concept compared to other programming languages like C and Java. One of the important features in Python is that it supports multiple programming paradigms, such as object-oriented, functional, and imperative programming or procedural styles (Srinath, 2017). The construct of Python programming language makes it easier for the user to write clear program codes on both a small and large scale.

The status of Python being among the fastest-growing programming languages is due to the sharp uptick in its use for data science. The board director of Python Software Foundation (PSF), Jacqueline Kazil, has predicted that Python's popularity will continue to grow, as the language's accessibility and utility continue to be attractive to researchers carrying out text analytics (Srinath, 2017).

2.1.6 Django Framework

Django (Holovaty & Kaplan-Moss, 2009) is a modern Python web framework that redefined web development in the Python world. The Django framework has its kick-start in around 2003, and it was a project done by Simon Willison and Adrian Holovaty at the Journal-World newspaper in Kansas, in the United States of America. It supports rapid and clean development processes, as well as pragmatic design.

Since Django is first built by experienced developers, it takes care of and pays more attention to the hassle of Web development. In other words, it understands the developer so well so that the developer can focus on writing the app without the need to reinvent the wheel (Shyam & Mukesh, 2020). A full-stack approach, pragmatic design, and superb documentation are some of the reasons for its success.

The framework initially uses a set of design principles that able to produce one of the most productive web development processes compared to many other web frameworks. It incorporates a set of design principles and trade-offs that make it one of the most productive frameworks for building the features needed by most medium to large web applications.

One of the strong points of using this framework is because Django can be considered as a full-stack framework. That means that its features cover everything from communication to databases, and from URL processing and to web page templating. Besides, Django projects, which are usually meant to be web sites are composed of a settings file, a URL mappings file, and a set of “apps” that provide the actual features of the web site.

2.1.7 Prototype

A prototype is a kind of like an early sample, model, or release of a product that is created to test a concept or process. Typically, a prototype is used to evaluate a new design to improve the accuracy of analysts and system users. The prototyping aims to establish accordance between the product requirements, ideas of designers, and users’ mental models. It is achieved by using usability testing or evaluation methods and tools. If the interface does not fully meet the needs of the target audience or system requirements, the change of the prototype will be made much easier than the final application.

People who are working in the development of technology are starting to realize more of the importance of the user interface design of software applications. Some experts believe that it should be started with “design, rather than just end with it. Design is an investment, not a cost“(Maeda, 2015). In software development, there are often used prototypes to receive feedback from users for refining the final product. (Neumann, 2004).

2.2 Related Works

In this work, the author presented the discussion of the development of SentiResearch as a web-based application system that implemented sentiment analysis. The author had done a

literature review before started the development process of the system. Based on the literature review that has been done, several pieces of research related to this work are found. The research found are about sentiment analysis web-based application.

According to one research (Sharma et al., 2018), sentiment analysis allows us to predict opinions and attitudes or groups of people or individuals. The researcher used twitter search API to retrieve live twitter data for any keyword required by the use and collected the tweets as the dataset to later be processed for the sentiment analysis. The research aimed to develop web-based application, that offered the visualization of sentiments associated with a hashtag, phrase or word on Twitter by plotting them on a map.

In another related research (Jain, 2017), the author of that research developed the web crawler to get the data from Twitter. The process is also part of opinion mining. It is believed that opinion mining will become a necessary and crucial part of the big companies and organizations and sentiment analysis result plays a big role for any major institutional change, product updates, reviews, launches, strategic planning and investments. The goal of it is to gain a competitive advantage in the modern era of internet and social platforms.

CHAPTER 3. ANALYSIS AND DESIGN

3.1 Research Framework

The idea of developing SentiResearch came from the author's own experience as a researcher in Information Technology (IT) company. This research began with the problem identification phase, continued to the literature review, and collect as many study resources as possible, then data collection with online review to validate the problem, to the development phase with Software Development Life Cycle methodology. At the end of the research, documentation is needed as part of the completion of the author's bachelor thesis. The research framework structured, as shown in Figure 3.1.

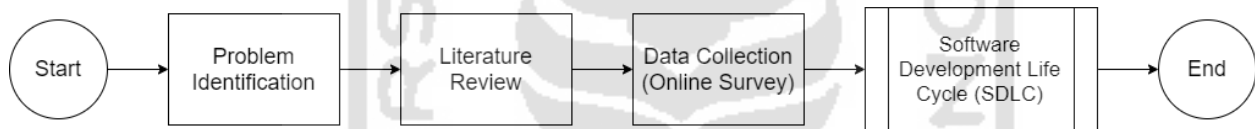


Figure 3.1 Research Framework

3.1.1 Problem Identification

Problem identification was done by conducting observation in the research team as the possible target user of this system, which are the people who work in a certain company, especially those whose role is a project manager or researcher. When conducting interviews, some questions have been asked to relevant parties to obtain information, system requirements, and assessment methods that are discussed by project manager, researcher, and other roles in the Engineering Product Development (EPD) department of certain companies in Indonesia. The results came as information, system requirements, and assessment methods that will be identified and can be supporting materials in the development of the system to be implemented.

3.1.2 Literature Study

A literature study was done by searching for some related research to strengthen the existing arguments and looking for data collection materials to develop the system. Based on the literature study, there were several references from previous studies related to the topic of this system development, such as the implementation of Django framework, Python programming language for text analytics, Sentiment Analysis, Data Preprocessing, and others. The purpose of the literature study is to obtain and have a deep understanding of the theoretical foundation needed in this study before moving on to the development process.

3.1.3 Data Collection

The method used in data collection is an online survey. The survey has been conducted before and in between the development of the system. It took almost a month until the completion. The survey has a total of 35 research respondents who are working in a research team from various prestigious IT companies like Google, Grab, Gojek, and many more. The data obtained including the roles and needs of Researcher in Qualitative Research Project. The aim of collecting data through online survey is to get better understanding of the work of Researcher in doing Qualitative Research project as well as to identify the needs and frustrations to be able to develop the solution.

The major advantage of doing an online survey is because it can be adjusted in a time-efficient manner. Online surveys also provide convenience in some other ways, for example, respondents can answer at any time when it is convenient for themselves; it is relatively simple for respondents to complete online surveys and for their responses to be analyzed by the author. Therefore, the set of questions in the online survey is attached below in Appendix A. Online Survey.

3.1.4 Software Development Life Cycle

The framework used in the development of SentiResearch is Software Development Life Cycle, specifically with the Prototype method. This framework, as seen in Figure 3.2 consists of detailed plans and tasks performed at each phase of the development process. The life cycle defines a methodology to improve the quality of software. The basic stages in the SDLC framework include planning, defining or analyzing, designing, building or developing, testing, and deployment.

In existence, there is an abundance of SLDC models designed and defined, which are followed during the software development process. Each process aims to ensure success in the process of software development. Among the most popular SDLC models are the Waterfall model, Iterative model, Spiral model, V-model, Big Band model, and Prototype model. The one that is implemented in the development of SentiResearch is the Prototype model.

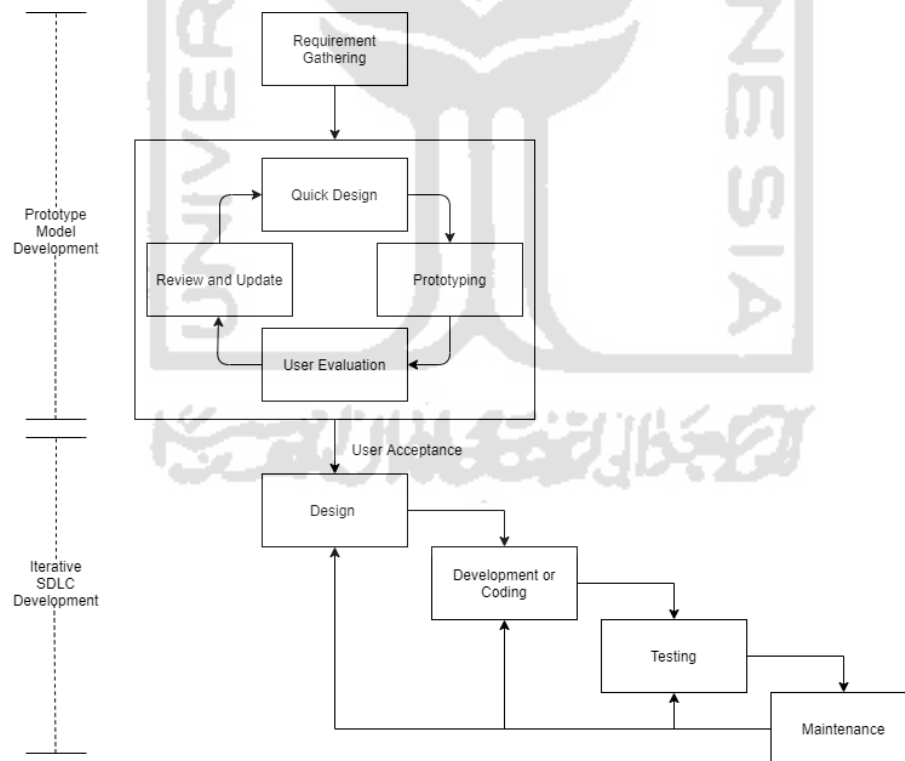


Figure 3.2 Prototype Method – Software Development Life Cycle

3.2 System Requirement Analysis

3.2.1 Actor Identification

An actor can be a person, organization, external system, or any role that has one or more interactions with the system. As the system will be used by the research team within a company, these roles in Table 3.1 are identified as the actors that will interact with the system:

Table 3.1 Actor Identification

Actor	Description
Administrator	The developer as the owner of this system. Responsible to manage the user and company registration.
Project Manager	The PM role in certain company, also main user of the system besides researcher. Responsible to register company and manage research project.
Researcher	The researcher role in certain company, also main user of the system. Responsible to manage research project from the designing, recruit respondents, to generating insight from sentiment analysis.
Other Employee	The other role in certain company. They are given the access to create account, join company workspace, then see the sentiment analysis result.

3.2.2 Functional Requirement

Functional Requirement (FR) determines and lists down a set of functions of the system. It consists of description of the functionality that the system has to offer. It lists down a set of components like the inputs and outputs to the software system, as well as the behaviors of certain actor. The

Table 3.2 shown below is the list of functional requirements of SentiResearch:



Table 3.2 Functional Requirement

Code	Description	Actor
FR-01	The application/system should support the user to register (signup) and sign in to the system before using	All Actors/Users
FR-02	The application/system should facilitate user to register a new workspace	All Actors/Users
FR-03	The application/system should facilitate user to join an existing workspace	All Actors/Users
FR-04	The application/system should facilitate the user to update or make an edit to their user profile data	All Actors/Users
FR-05	The application/system should facilitate the user to manage (create new, edit, delete, and view details) qualitative research project data in the system	Researcher, Project Manager
FR-06	The application/system should facilitate the user to manage (create new, edit, delete, and view details) research respondent data in the system	Researcher
FR-07	The application/system should facilitate the user to manage (create new, edit, delete, and view details) research question data in the system	Researcher
FR-08	The application/system should facilitate the user to manage (create new, edit, delete, and view details) research answer data in the system	Researcher
FR-09	The application/system should facilitate the user to generate sentiment analysis for the research result from interview narrative (answer) in the	All Actors/Users

	system	
--	--------	--

3.2.3 Non-Functional Requirement

Non-Functional Requirement (FR) specifies the system's operation capabilities and constraints that enhance the functionality of the system. The types of non-functional requirement include security, reliability, scalability, user experience, and many more. The Table 3.3 shown below is the list of non-functional requirements of SentiResearch:

Table 3.3 Non-Functional Requirement

Code	Description
NFR-01	The application/system should have a suitable user interface which is considered to be minimalist, with simple data entry form so that filling out forms can be done quickly.
NFR-02	The application/system should be able to limit the access rights of each Actor by the authentication process and registration before using the system.
NFR-03	The application/system should be able to present the data in the form of a clear dashboard with a graphical interface that is easy to understand.

3.3 System Design

Systems design is the phase of defining and designing the elements of the system such as the architecture, data, modules, and interfaces to satisfy specified system requirements. This phase came after the requirement has been analyzed. The design of SentiResearch system put

emphasize on the site mapping, database design, Django architecture, and low-fidelity prototyping.

3.3.1 Use Case Diagram

The use case diagram shown below represents a user's interaction with the system, that has been listed down on functional requirement table. It models the system's functionality using actors and use cases. The use case diagram can be seen in Figure 3.3.

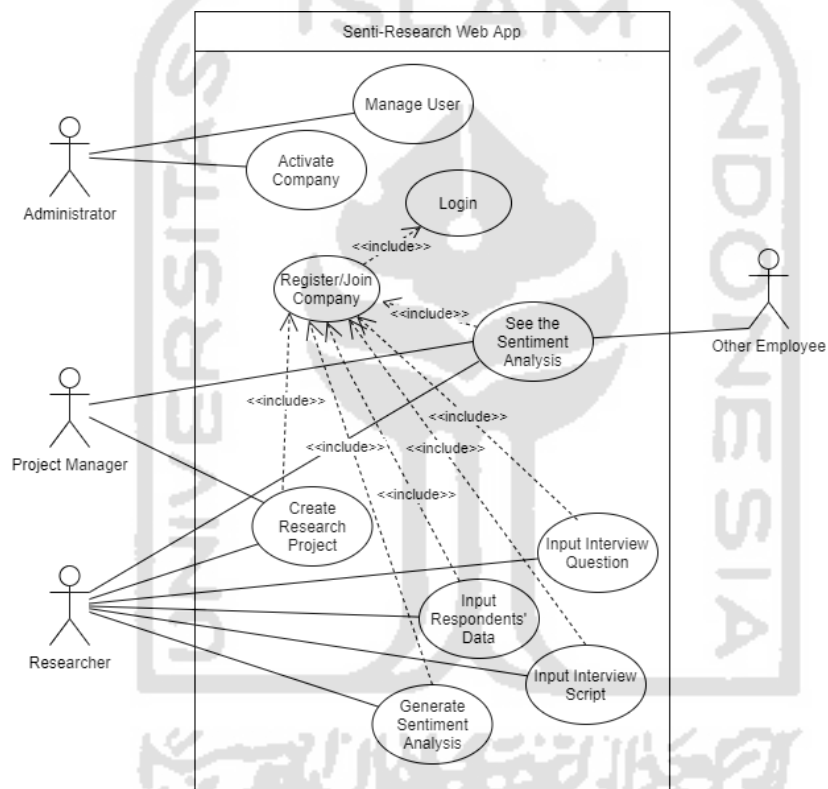


Figure 3.3 Use Case Diagram

3.3.2 Site Map

A sitemap is a mapping file where the site provides information about the pages, media, and files and the relationships or user actions between them. It is commonly used by website designers to plan the pages of a website that is going to be developed. The figures below are the sitemap of SentiResearch.

1. Admin

The administrator of SentiResearch can log in to the specific site provided by the Django framework. The automatic admin site is considered to be one of the most popular advantages of Django. The data that can be managed by the admin is customizable. Therefore, the SentiResearch system allows admin to be able to manage only two data models, which are user and company, as shown in Figure 3.4.

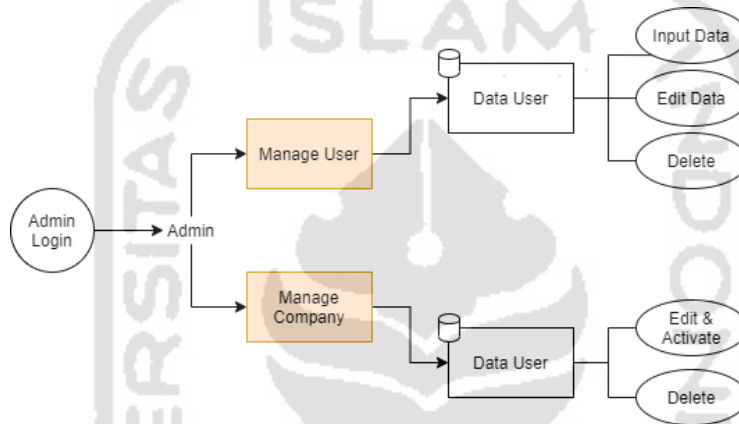


Figure 3.4 Admin Sitemap

2. Project Manager

The user that has set the role as a project manager can only manage the research project to assign it to a specific researcher and also generate the sentiment analysis of it. The site map of the project manager can be seen in Figure 3.5.

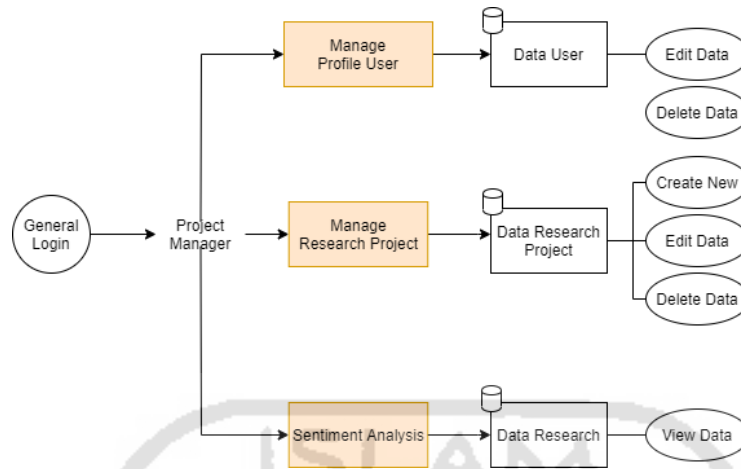


Figure 3.5 Project Manager Sitemap

3. Researcher

The user that has set the role as a researcher can get the most access in SentiResearch, as shown in Figure 3.6. The researcher can manage almost every data model in the system, including a research project, respondent, question, answer, and then generate its sentiment analysis.

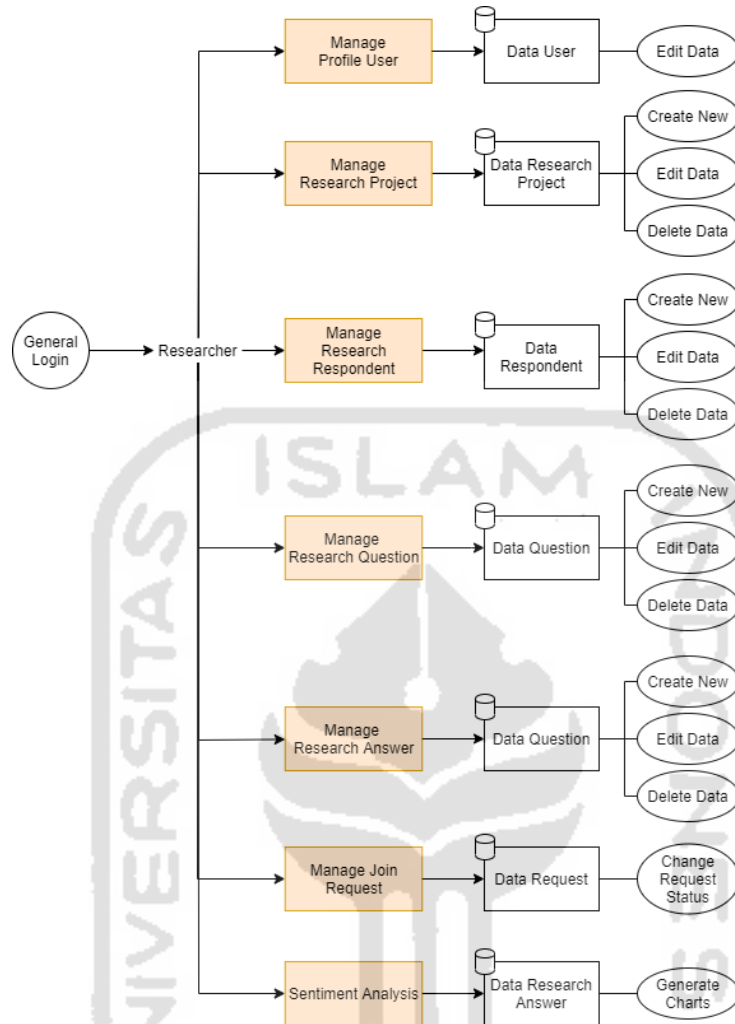


Figure 3.6 Researcher Sitemap

4. General User

The other role of the user can only get to access the sentiment analysis generator besides to also be able to edit the profile just like other users. The site map of a general user can be seen in Figure 3.7.

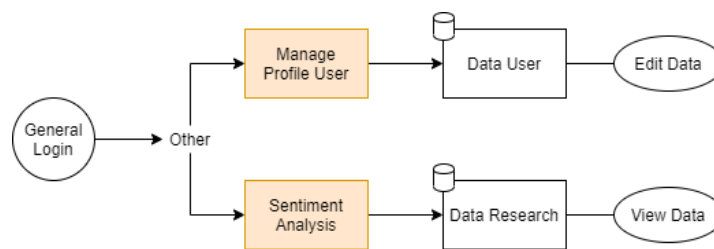


Figure 3.7 General User Sitemap

3.3.3 Database Schema Design

The following Figure 3.8 shows the database schema design that is generated from MySQL Workbench as the visual database design tool. A relational database is chosen because it allows sorting based on any field and generates reports that contain only certain fields from each record.

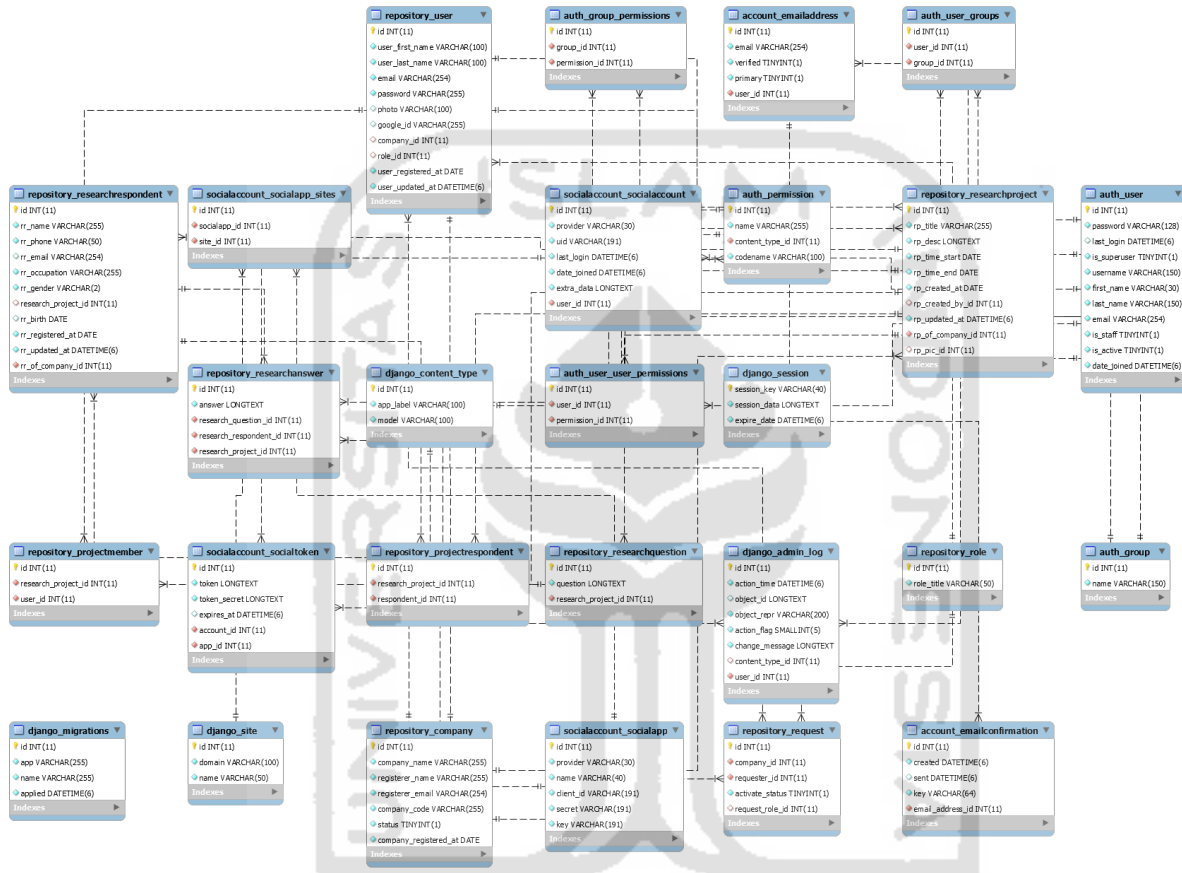


Figure 3.8 Database Schema

3.3.4 Architectural Design

The architecture follows the Model-View-Template (MVT) architecture provided by the Django framework. A web-framework usually has a basic Model-View-Controller (MVC) architecture software design pattern for developing social web applications. Django works differently with MVT, and it is slightly different from the common MVC. The architectural design is presented in Figure 3.9.

1. Model, a model is like a data handler between database and view. It is the logical data structure behind the entire application and is responsible for maintaining the data. Model is represented by a database (e.g., SQL, MySQL, PostgreSQL, MariaDB, etc.). The model is in accordance with the Database Schema Design.
2. View, a view in this MTV architecture is responsible for formatting the data via the model. In turn, it transfers the data model to the template for viewing. More detailed views are discussed in the URL Configurations section.
3. Template, a template consists of static parts of the desired HTML as the output and some syntax to describe the dynamic content of the application.

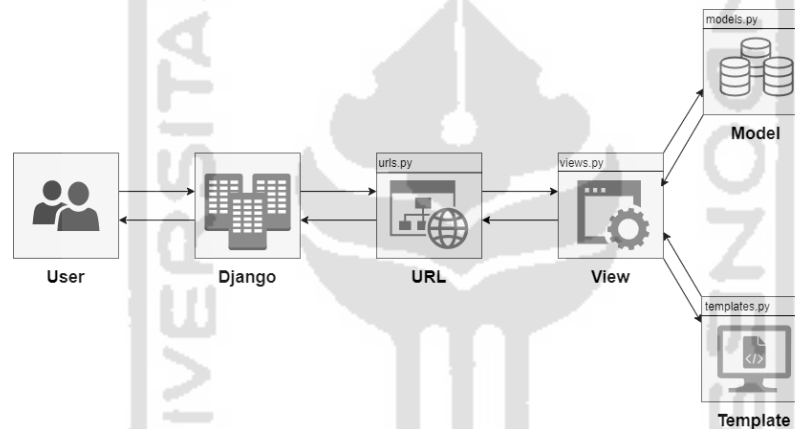


Figure 3.9 Architecture Design

3.3.5 System Low Fidelity Prototype

A prototype is an early sample of the final product and simulation of the interaction between the user and the designed interface. The purpose of modeling a prototype is to test the usability and feasibility of the interface designs (Babich, 2017). A prototype has different types of fidelity which are low and high fidelity. The fidelity of the prototype is decided based on the goals of prototyping, completeness of design, and available resources. In this research, the low fidelity prototype has been created with the Balsamiq tool and tested to the target user in order to get insights and feedbacks until the design concept works as intended.

1. Landing Page Mockup

The landing page is designed, as shown in Figure 3.10. The design is rather simple, with no data action. The purpose is to give the user an overview of the system.

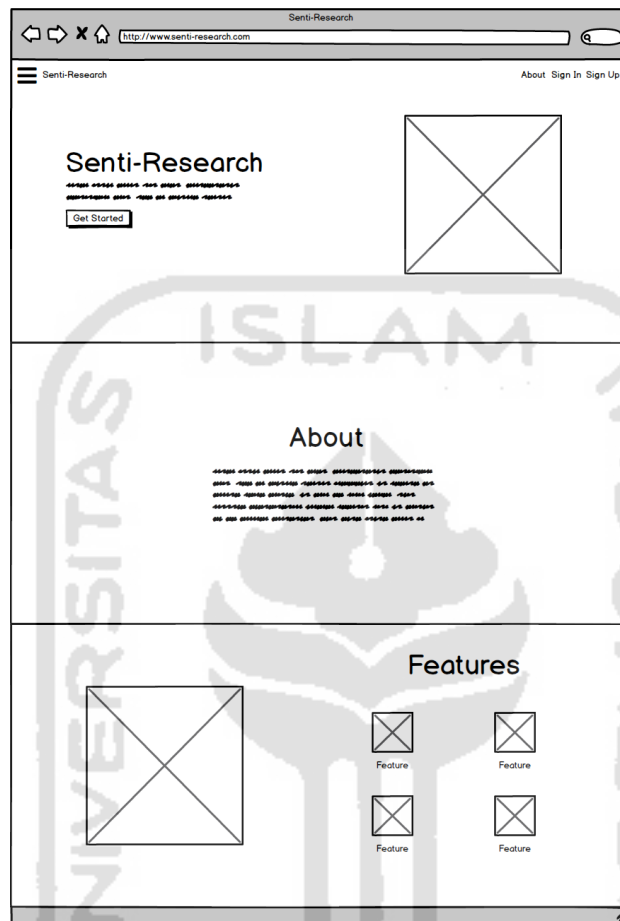


Figure 3.10 Landing Page Mockup

2. Authentication Mockup

The authentication consists of sign up and sign in before accessing the system and is pictured in Figure 3.11.

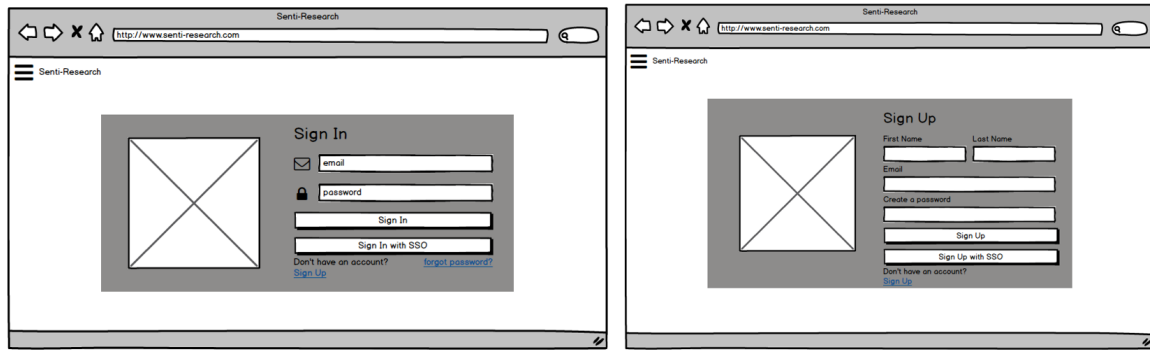


Figure 3.11 Authentication Mockup

3. Register and Join Workspace Mockup

Register and join used to get access to the workspace. After signing in, the user can choose to create and register a new workspace for their company/organization or join the existing one. The mockup is shown in Figure 3.12.

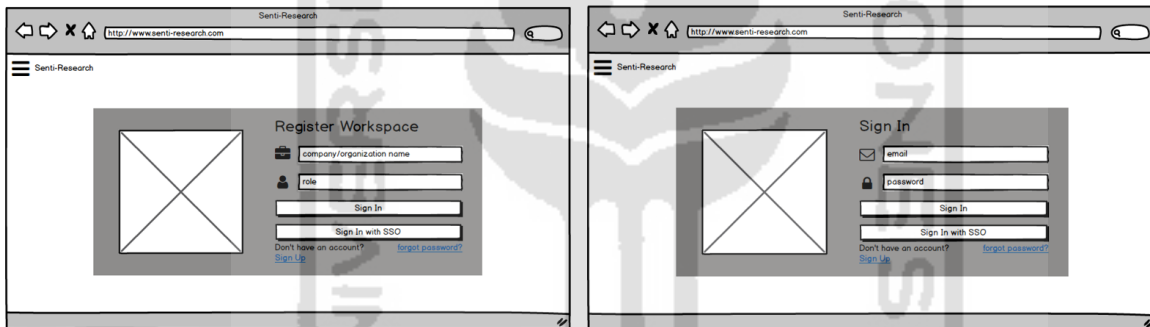


Figure 3.12 Register and Join Workspace Mockup

4. Workspace Dashboard Mockup

The workspace dashboard visualizes the data with various of the chart. This page displays the overview of crucial data saved in the system as data and charts are the essential things for a research team. The mockup is shown in Figure 3.13.

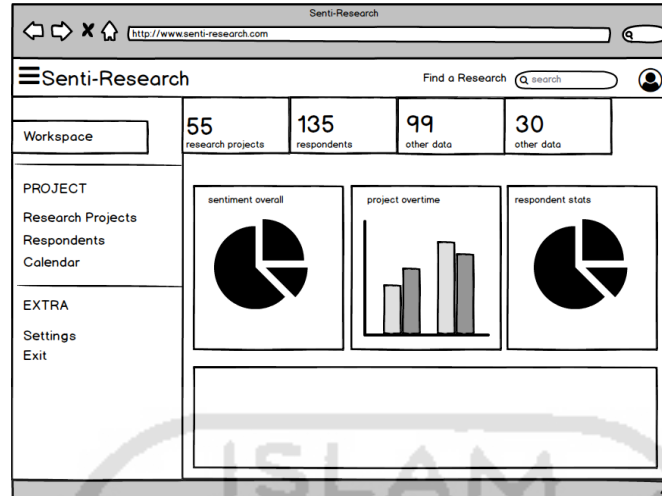


Figure 3.13 Workspace Dashboard Mockup

5. Research Details Mockup

The following Figure 3.14 shows the detailed information of certain research projects.

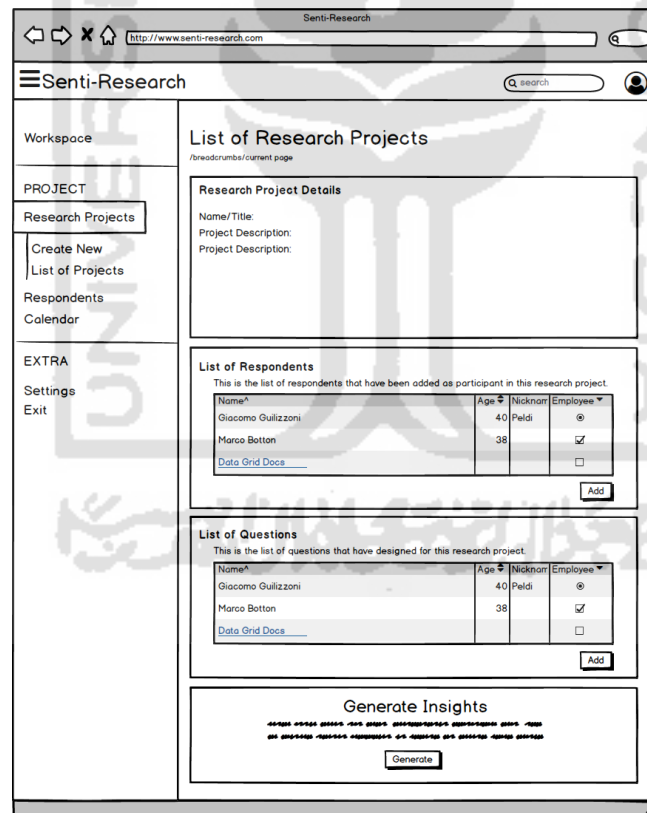
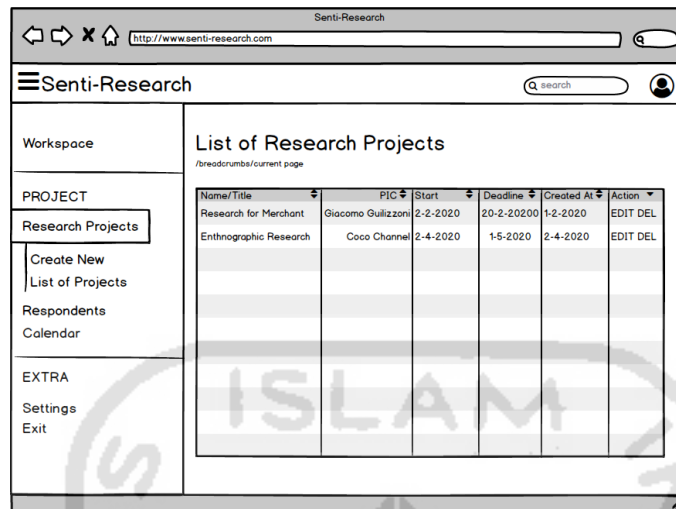


Figure 3.14 Research Details Mockup

6. Research Table Mockup

The list of data in a certain table is displayed in table mockup. The following Figure 3.15

shows the example of a research table that stores research project data.

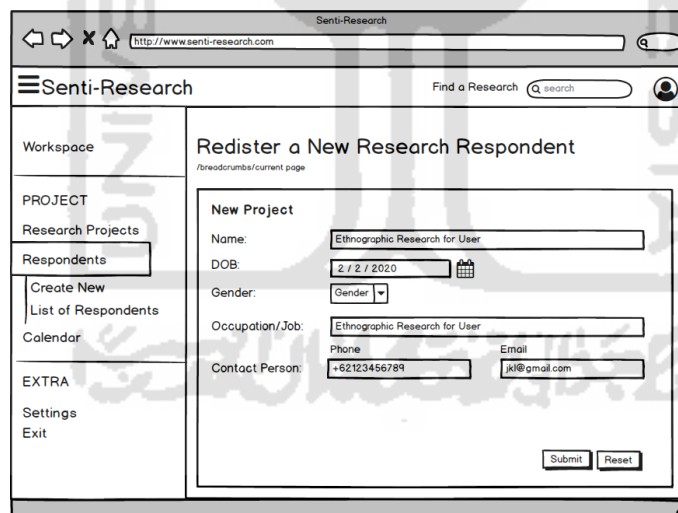


Name/Title	PIC	Start	Deadline	Created At	Action
Research for Merchant	Giacomo Gulizzoni	2-2-2020	20-2-2020	1-2-2020	EDIT DEL
Ethnographic Research	Coco Channel	2-4-2020	1-5-2020	2-4-2020	EDIT DEL

Figure 3.15 Research Table Mockup

7. Research Form Mockup

The following Figure 3.16 shows the example of form page. The user will have to input the information needed into this kind of form in the system.



Redister a New Research Respondent

New Project

Name:

DOB:

Gender:

Occupation/Job:

Contact Person:

Figure 3.16 Research Form Mockup

8. Sentiment Analysis Insight Chart Mockup

The main feature of this system is the sentiment analysis tool that allows the user to generate insight charts needed for research reporting. There are several types of charts

used to visualize the data, as shown in Figure 3.17.

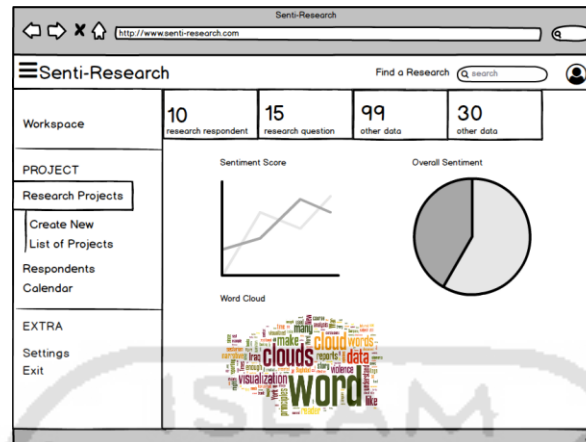


Figure 3.17 Sentiment Analysis Insight Chart Mockup

3.4 Sentiment Analysis Methodology

In the field of computational linguistics as well as computer science, the study and research on sentiment analysis have been gaining more popularity in natural language processing, information retrieval, and data mining. There is a variety of methodologies used in sentiment analysis, like machine learning, lexicon, and hybrid. Despite a large number of researches that have been done with machine learning-based methodology, the author chose to implement lexicon-based instead. This is due to limited data resources to process the data training phases needed in machine learning-based. The methodology used for sentiment analysis in this research is the lexicon dictionary-based sentiment analysis. This methodology is chosen due to the limited data.

The research of lexicon dictionary-based sentiment analysis mostly focused on various construction of lexicon in the English language, such as AFINN Lexicon (Nielsen, 2011), SentiWordNet (Baccianella et al., 2010), Liu Lexicon (Liu et al., 2005). The number of researches that focuses on the Indonesian language is very limited, causing many researchers to translate the popular lexicon from English to other languages, including Bahasa Indonesia.

The set of phases pictured by the flowchart shown below in Figure 3.18, those phases define how the sentiment analysis works. The process started with text preprocessing to get the

desirable dataset, to assign polarity score using InSet lexicon data, then calculate the score before the data charts can be generated for the user to evaluate and analyze the sentiment result.

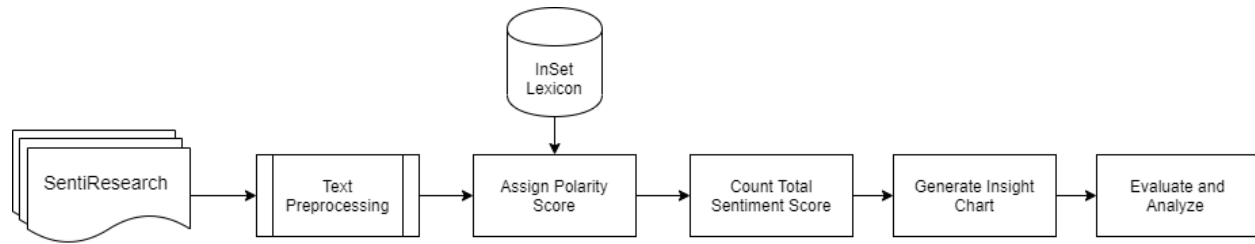


Figure 3.18 Sentiment Analysis Flowchart

3.4.1 Text Preprocessing

Text preprocessing is a process of changing the form of unstructured data into structured data as needed for further mining processes (sentiment analysis, summarizing, document clustering, etc.). It converts the text in sentence or document into a form that is analyzable and predictable for text analytics tasks. There are different ways and methods to preprocess the text in sentences or documents. The preprocessing methods needed for this research can be seen in Figure 3.19.

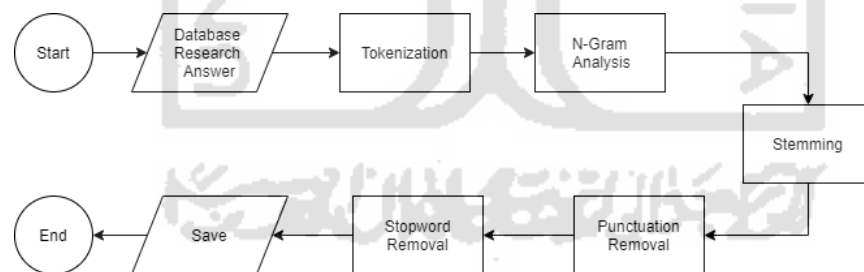


Figure 3.19 Text Processing Flowchart

1. Tokenization

Tokenization is the most commonly used text preprocessing step. It works by breaking a stream of text into words, symbols, phrases, or other meaningful elements of text called tokens. The aim of the tokenization is the exploration of the words in a sentence. The list

of tokens becomes the input for processing, such as text mining or parsing (Gurusamy & Kannan, 2016). SentiResearch implemented the Natural Language Toolkit (NLTK) library as the tokenizer.

2. N-Gram Analysis

N-Gram is a sequence of n items retrieved from a text or speech. The items can be syllables, letters, or words depending on the application. Typically, the n-grams can be collected from a text or speech corpus (large body of text). In this case, the n-gram analysis works with a tokenizer to find a word that contains a hyphen.

3. Stemming

Stemming is the process of chopping off the ends of words, and it often includes the removal of derivational affixes. The goal of stemming is to reduce inflectional forms and sometimes derivationally related forms of a word to a common base form. SentiResearch implemented the Sastrawi library as the stemmer to reduce inflected words in Bahasa Indonesia (Indonesian Language) to their base form.

4. Punctuation Removal

This process deletes all punctuations from the text. Punctuation itself is a set of conventional signs to make it easier for reading and understanding of the written text. The set of punctuation that needs to be removed includes !()-[]{};:'",<>./?@#\$\$%^&*~.

5. Stopword Removal

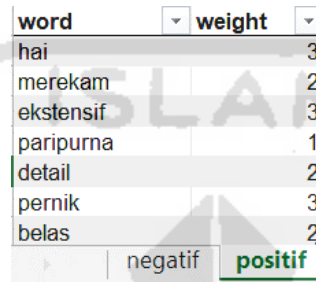
A stop word list is a list of commonly repeated features that emerge in every text document. The common features such as conjunctions such as 'or,' 'and,' 'but' and pronouns 'he,' 'she,' 'it' etc. need to be removed due to it does not have an effect, and these words add a very little or no value on the classification process (Kadhim, 2018).

3.4.2 InSet Lexicon Dictionary-Based

InSet lexicon is an Indonesian lexicon containing 3,609 positive words and 6,609 negative words with a range of -5 to +5. The dictionary of sentiment lexicon utilized the *Twitter* dataset

that has been filtered with Bahasa Indonesia with the result of two kinds of emoticons that categorized as positive and negative polarity.

The author has set the file extension of the InSet lexicon dictionary to be .xlsx and has divided the type of sentiment into two different sheets in the file. Then the file is imported to the system by using a certain python library. The example of the word and its weighting list in a lexicon dictionary dataset can be seen in Figure 3.20.



word	weight
hai	3
merekam	2
ekstensif	3
paripurna	1
detail	2
pernik	3
belas	2

negatif **positif**

Figure 3.20 InSet Lexicon Dictionary in Excel

CHAPTER 4. IMPLEMENTATION AND TESTING

4.1 System Development Specification

In developing this system, there are system development specifications that are used to simplify the system development process; the system development specification details can be seen in Table 4.1.

Table 4.1 System Development Specification

Aspect	Specification
Appilcation Base	Web Application
Platform	Python V.3.7.6
Framework	Django V.3.0.5
Database	MySQL
Browser	Firefox, Chrome, Microsoft Edge

Operation System	Windows 10 64-bit
IDE	Visual Studio Code
Processor	Intel® Core™ i7-6700HQ
RAM	16GB

4.2 Back-End Implementation

4.2.1 Installed Modules and Packages

Python modules and packages are two components that facilitate the process of breaking a large task into smaller, separate, and more manageable subtasks or modules; it is called a modular programming. The installed modules and packages can be seen in Table 4.2.

Table 4.2 Modules and Packages

Modules and Packages	Description
appdirs==1.4.3	This module determines appropriate platform-specific dirs so that other Python packages can include their private copy.
distlib==0.3.0	Low-level components of distutils2/packaging to make packaging in Python easier.
django-allauth==0.41.0	A set of Django applications that addresses authentication, registration, account management, and even third party (social) account authentication
importlib-metadata==1.5.0	A library to get access to the metadata for a Python package.
nltk==3.5	The Natural Language Toolkit (NLTK) works as a Python package for Natural Language Processing (NLP).
oauthlib==3.1.0	It implements the logic of OAuth1 or OAuth2 without assuming a

Modules and Packages	Description
	web framework or specific HTTP request object.
openpyxl==3.0.3	A Python library to write or read Excel files (xlsx/xlsm/xltx/xltn).
pylint==2.4.4	A Python static code analysis tool that inspects any errors, helps enforcing a coding standard, and offers refactoring suggestions.
regex==2020.5.7	This module helps with case-insensitive matches in Unicode, nested sets and set operations, flags, any many more.
requests==2.23.0	A HTTP library that enables sending HTTP/1.1 requests easier.
requests-oauthlib==1.3.0	It uses the Python Requests and OAuthlib libraries to provide an easy-to-use Python interface for building OAuth1 and OAuth2 clients.
Sastrawi==1.0.1	A Python library that helps with stemming process in NLP; it reduces inflected words in Bahasa Indonesia to their base form.
virtualenv==20.0.10	A Python package to create and manage an isolated environment for Python projects.

4.2.2 URL Configurations

URL Configuration works like a table of content in Django-powered web application. This model is mapping between URL paths to Python view functions that should be called in those URLs. The view functions are defined in the file named views.py and then collected all together with the URLs in the urls.py file. In this system, some Django applications has been created mainly in order to make the project structure well-organized. There are landing, repository, and semantic applications; each consists of views.py and urls.py file to make the dynamic web application.

1. Landing App urls.py

The landing application basically work as the frontier separating the static landing page and the workspace by providing authentication, registration, and access management. The URLs and views of landing app are showed in Table 4.3

Table 4.3 Landing App URLconf

URL	View	Description
path("")	index	Show landing page (home) of SentiResearch
path('signin/')	signin	Sign into the workspace
path('signup/')	signup	Create an account and sign up
path('company-register/')	register	Register new workspace for company/organization
path('company-join/')	join	Join existing workspace of company/organization

2. Repository urls.py

The repository application allows user to use the main feature which is the workspace. It consists of the basic Create Read Update Delete (CRUD) functions and most template resources that are displayed in the URLs. The URLs and views of repository app are showed in Table 4.4.

Table 4.4 Repository App URLconf

URL	View	Description
path('workspace/')	workspace	Show workspace dashboard
path('profile/')	profileview	Show data user in user profile page
path('profile-edit/')	profileedit	Edit user data in profile page
path('request-join/')	requestjoin	Manage (confirm) request join
path('workspace/research-brief/')	researchbrief	Show research brief page
path('workspace/research-intro/')	researchintro	Show research intro page

URL	View	Description
path('workspace/research-new/')	researchnew	Input form for new research project
path('workspace/research-list/')	researchlist	Show data table of research project
path('workspace/<str:pk>/research-view/')	researchview	Show research project detailed data and manage project respondent
path('workspace/<str:pk>/research-edit/')	researchedit	Edit research project data
path('workspace/<str:pk>/research-delete/')	researchdelete	Delete research project data
path('workspace/respondent-new/')	respondentnew	Input form for new research respondent
path('workspace/respondent-list/')	respondentlist	Show data table of research respondent
path('workspace/<str:pk>/respondent-view/')	respondentview	Show research respondent detailed data
path('workspace/<str:pk>/respondent-edit/')	respondentedit	Edit research respondent data
path('workspace/<str:pk>/respondent-delete/')	respondentdelete	Delete research respondent data
path('workspace/<str:pk>/projectrespondent-delete/')	rrespondentdelete	Delete project respondent data
path('workspace/<str:pk>/question-add/')	questionadd	Input form to add research question
path('workspace/<str:pk>/question-delete/')	questiondelete	Delete research question data

URL	View	Description
-delete/')		
path('workspace/<str:pk>/<str:respondent>/answer-add/')	answeradd	Input form to add research answer
path('workspace/<str:pk>/<str:respondent>/answer-view/')	answerview	Show the detailed data of questions and answers
path('workspace/detail/')	workspacedetail	Show the detailed data of the workspace
path('logout/')	logout	To logout from the workspace

3. Sentiment urls.py

The sentiment application has the least mapping between URLs and views. It only has one which works to generate chart insight from the sentiment analysis phase. The URLs and views of sentiment app are showed in Table 4.5.

Table 4.5 Sentiment App URLconf

URL	View	Description
path('generate/<str:pk>/')	generate_sentimen	Generate sentiment analysis as well as the insight chart

4.3 Front-End Implementation

Django is a framework that originally has lots of files and folders inside it. A well-organized project really helps the administrators and developers to find the proper path of files and folder easily. Therefore, the front-end implementation of this system follows Django project structure as displayed in Figure 4.1.

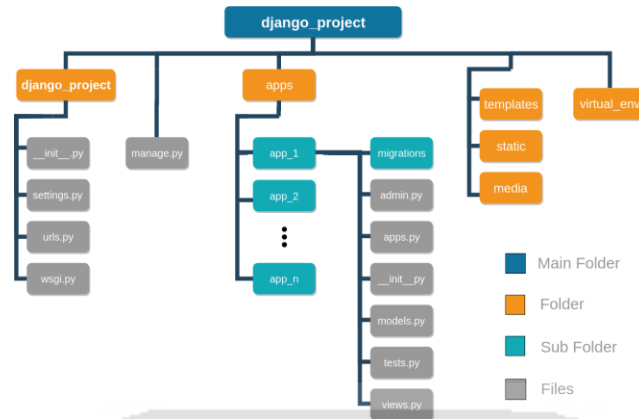


Figure 4.1 Django Project Structure

Django itself is a collection of applications, each designed to work on one big thing. In the development of this system, one project named 'senti_research' is created as well as three other applications which are the 'landing', 'repository', and 'sentiment'. The front-end implementation is explained further as follows:

4.3.1 Landing App User Interface

1. SentiResearch Landing Page

The Figure 4.2 shown below is the landing page of SentiResearch web application. This is the static templates to welcome the user. The page consists of information that describes the system with the about and feature section.

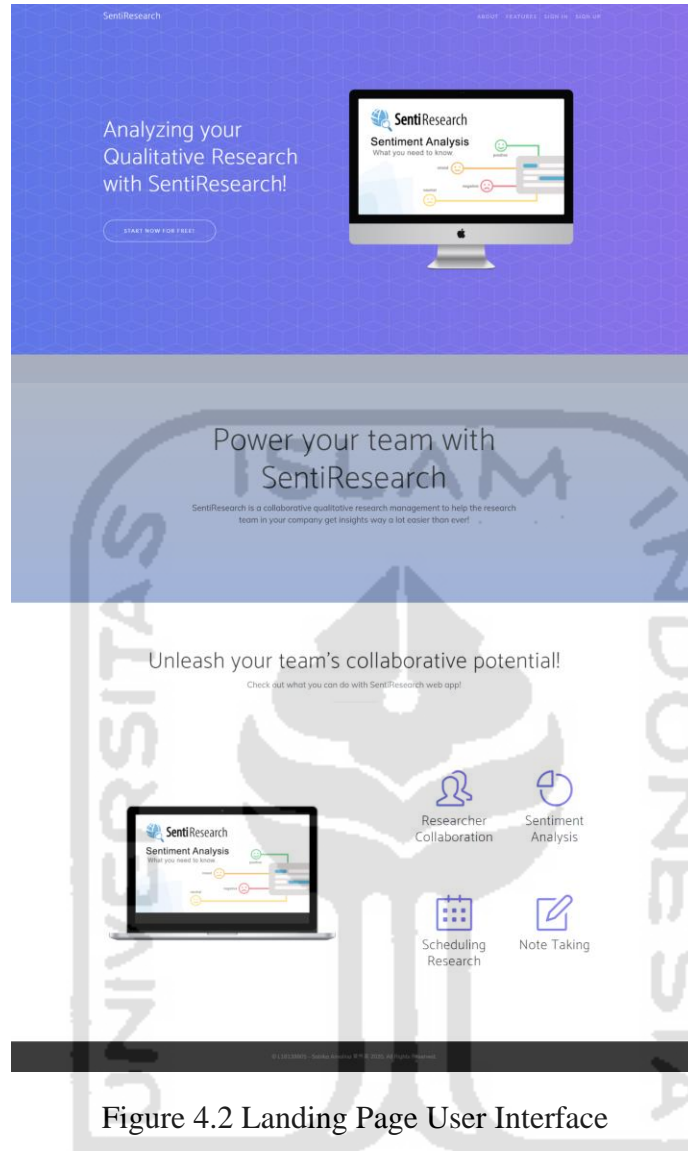


Figure 4.2 Landing Page User Interface

2. Sign in Page

The Figure 4.3 shown below is the sign in page where user with registered account can log into the system and the redirect to the workspace page. The user can sign in with the registered email address and password.

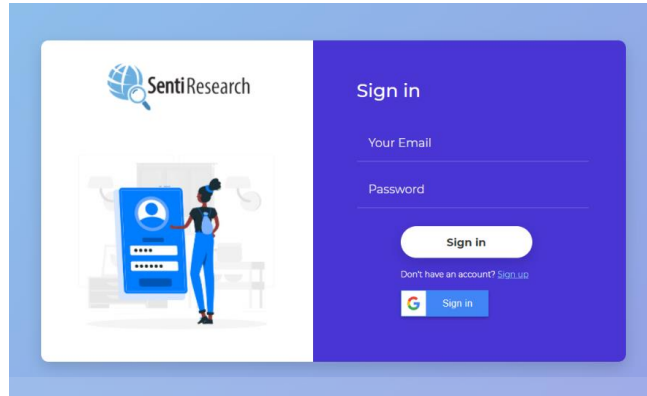


Figure 4.3 Sign in Page User Interface

3. Sign up Page

The Figure 4.4 shown below is the sign up page where new user creates an account. The user has to fill in first and last name, email, and create a password for the account. The new user has not been assigned to any workspace so after the account is created successfully, the user will have to register or join to a workspace.

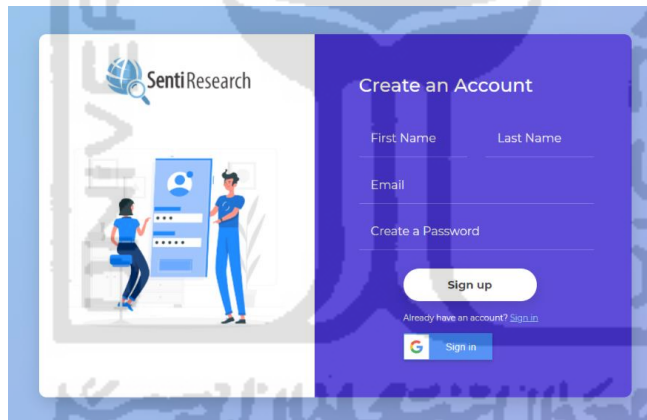


Figure 4.4 Sign up Page User Interface

4. Workspace Register Page

The Figure 4.5 shown below is the workspace register page where the user create new workspace for the company or organization. Some information like the company name and job title are needed in this registration form. The role is used to determine the access; the user has to be a researcher or project manager to register a new workspace.

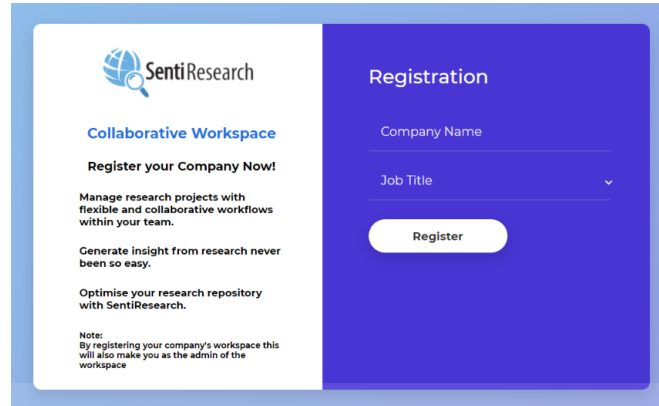


Figure 4.5 Workspace Register Page User Interface

5. Workspace Join Request Page

The Figure 4.6 shown below is the workspace join page where the user joins an existing workspace that has been registered before by other users. The user has to input the company code that is generated randomly by the system. The only one that has access to see the code is the workspace registerer.

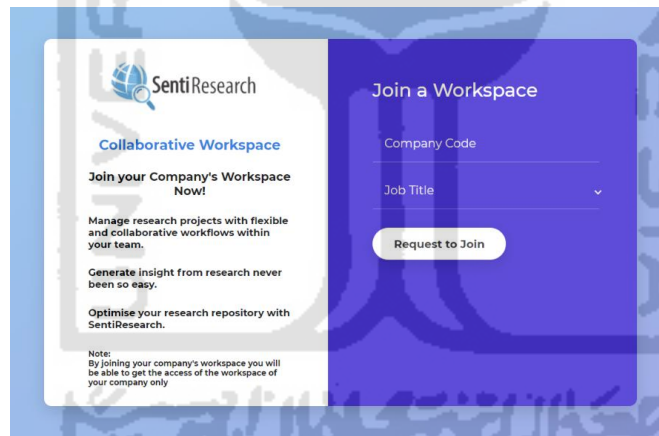


Figure 4.6 Workspace Join Request Page User Interface

4.3.2 Repository App User Interface

1. Workspace Dashboard Page

The Figure 4.7 shown below is the workspace dashboard as the first page the user will be directed to after sign in. This dashboard visualizes the data in the workspace with a variety of charts. The research team is commonly in need of chart to present the research

data results to the others so that is why the dashboard plays a big role in this system.

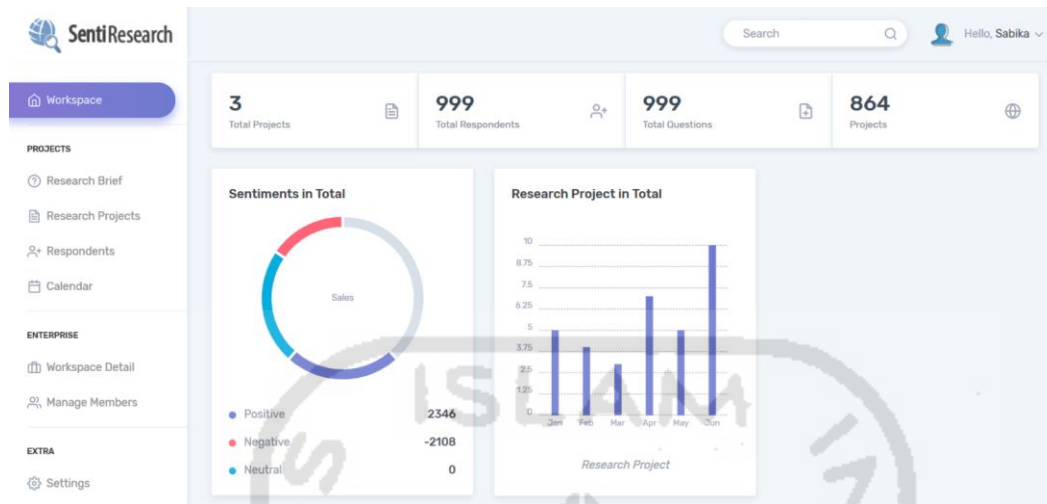


Figure 4.7 Workspace Dashboard Page

2. Research Project Management Page

The Figure 4.8 shown below is the research project management page where all of the research projects in user's workspace are listed down with crucial information like the name/title, the person in charge, and the project schedule. The user can manage by editing or deleting the research project from the data.

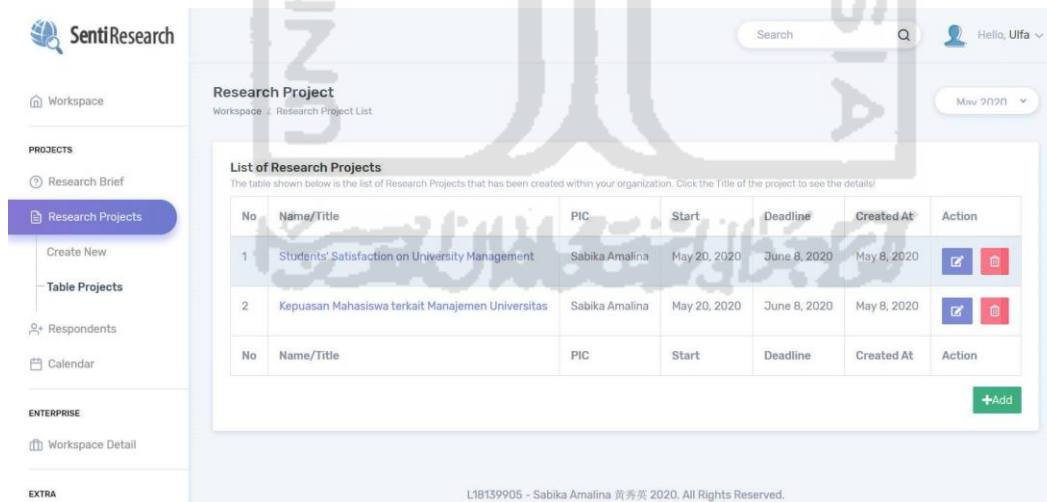


Figure 4.8 Research Project Management Page User Interface

3. Research Project Details Page

The Figure 4.9 shown below is the research project details page that displays the detailed

information of the research project. The other project members can read the details thoroughly before start working on the research project.

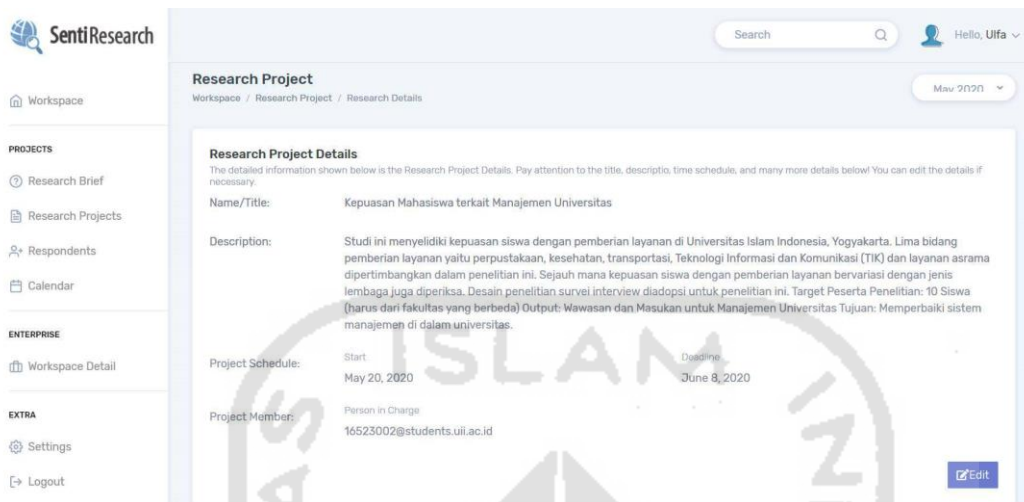


Figure 4.9 Research Project Details Page User Interface

4. Project's Respondent Management Section

The Figure 4.10 shown below is the project's respondent management section within the research project details page. In this section, the user can manage the project's respondent data by clicking the delete button and edit to redirect to the edit form page. The user can also add new project's respondents by filling the name that has been registered.

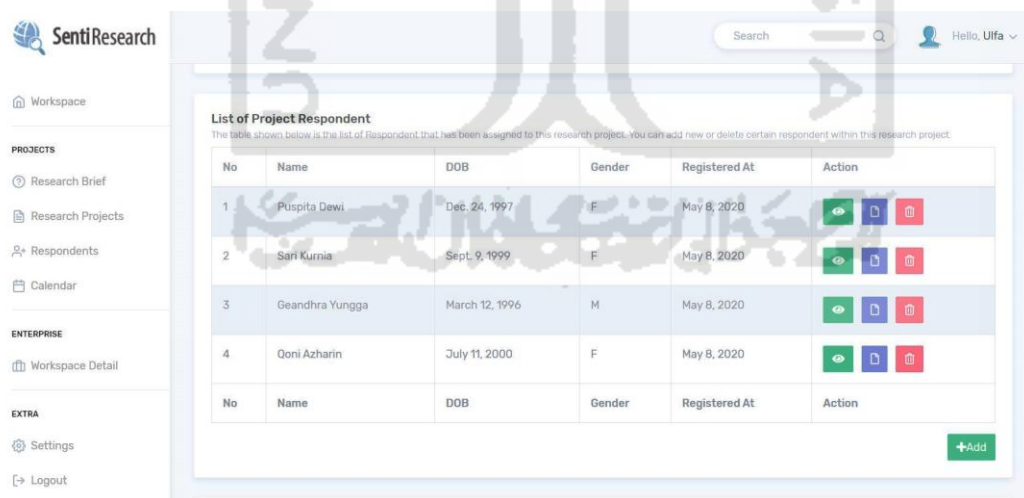


Figure 4.10 Project's Respondent Management Section User Interface

5. Project's Question Management Section

The shown below Figure 4.11 is the is the project's question management section within

the research project details page. In this section, the user can add questions that will be asked in the interview to the system.

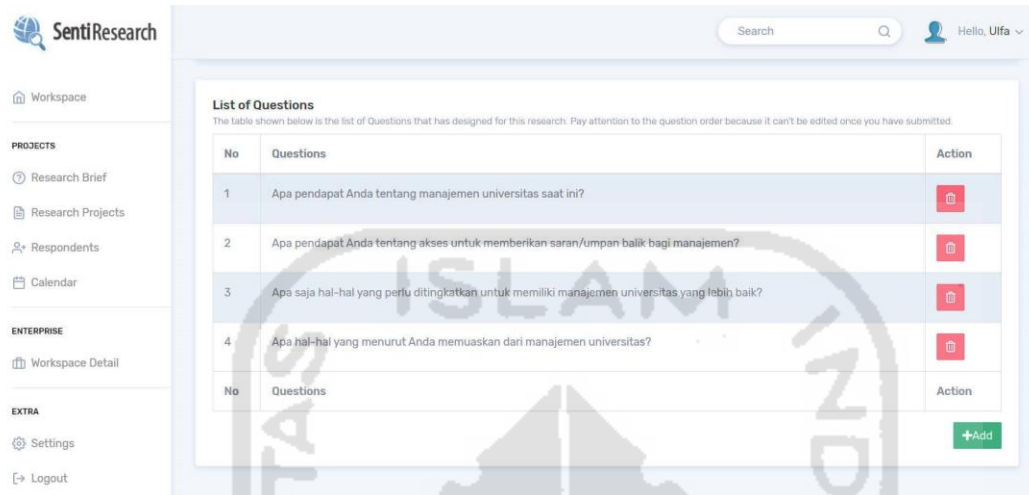


Figure 4.11 Project's Question Management Section User Interface

6. Project's Question and Answer Page

The shown Figure 4.12 below is the project's question and answer page. This page displays the detailed information of the research project as well as the question and answer.

Figure 4.12 Project's Question and Answer Page User Interface

7. Research Respondent Management Page

The shown Figure 4.13 below is the research respondent management page where the user register and create new research respondent data. This process is needed before the user can add and assign the respondent to certain research project.

No	Name	DOB	Gender	Registered At	Action
1	Puspita Dewi	Dec. 24, 1997	F	May 8, 2020	
2	Sari Kurnia	Sept. 9, 1999	F	May 8, 2020	
3	Goni Azharin	July 11, 2000	F	May 8, 2020	
4	Geandhra Yungga	March 12, 1996	M	May 8, 2020	
5	Zainuddin Zukhri	Feb. 4, 1978	M	May 8, 2020	

Figure 4.13 Research Respondent Management Page User Interface

8. Research Respondent Details Page

The shown Figure 4.14 below is the research respondent details page. The information about research respondent that has been recorded in the system can be seen in this page. Not all roles of user can get the access to this page because the respondent's information is considered sensitive and part of the privacy.

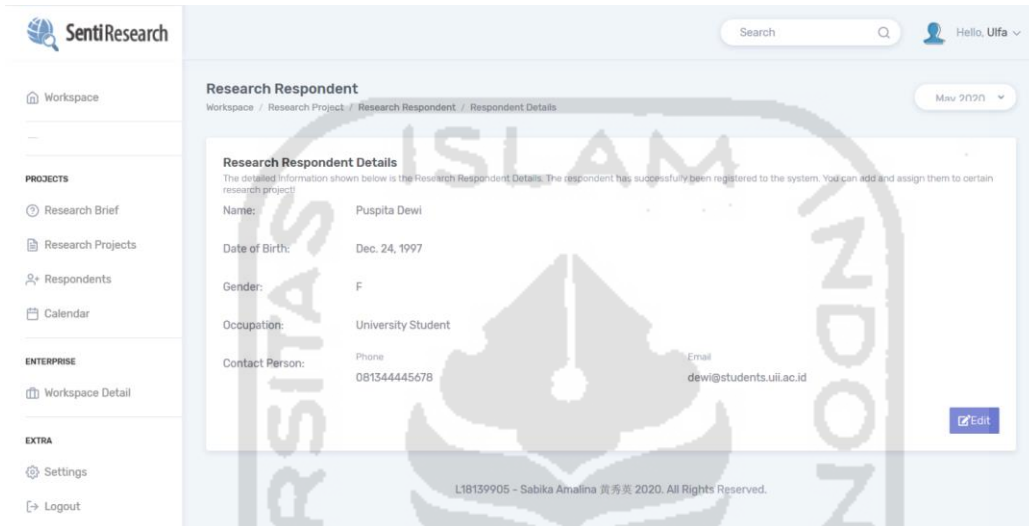


Figure 4.14 Research Respondent Details Page User Interface

9. Workspace Details Page

The shown Figure 4.15 below is the workspace details page that shows the detailed information regarding of the workspace. The information displayed such as the workspace name, the user as registerer, and code. The code is randomly generated by the system and is needed for the user to join an existing workspace.

Company/Organization Information

Workspace / Workspace Details

Workspace Details
The detailed information shown below is the Company/Organization Workspace Details. The company has successfully been registered to the system by the registerer. You can see the details below!

Name: Universitas Islam Indonesia
Code: 45716E
Registerer: Name: Sabika Amalina, Email: 16523002@students.uii.ac.id
Registered: May 8, 2020

List of Workspace Members
The table shown below is the list of Users/Members that has been registered to this Workspace.

No	Name	Email	Role
1	Sabika Amalina	16523002@students.uii.ac.id	Project Manager
2	Ulfa Amalia	16523151@students.uii.ac.id	Researcher

L18139905 - Sabika Amalina 黃秀英 2020. All Rights Reserved.

Figure 4.15 Workspace Details Page User Interface

10. Workspace Join Request Management Page

The shown Figure 4.16 below is the join request management page that list down the users who request to join and the chosen role. Only the registerer of the workspace can manage this join request data.

Workspace Join Member Request

Workspace / Research Project / List of Join Requests

List of Join Requests

No	Name	Email	Role	Action
1	Ulfa Amalia	16523151@students.uii.ac.id	Researcher	Deactivated
2	Fionna Saphira	16523018@students.uii.ac.id	Researcher	Deactivated

L18139905 - Sabika Amalina 黃秀英 2020. All Rights Reserved.

Figure 4.16 Workspace Join Request Page User Interface


```
from django.contrib.staticfiles.storage
import staticfiles_storage
from repository import models
from openpyxl import Workbook, load_workbook
import os
```

Figure 4.18 Modules and Packages Imported

As shown in Figure 4.18, there are modules and packages that should be imported. The data that used in the process of text preprocessing as well as the sentiment scoring, and classification was the internal data of the system. To get access to the data, the model of where the data belongs to should be imported to the project. In this project, the research answer was the data that the author used to get the sentiment and it is located in the repository app model.

Other modules needed are the staticfiles_storage as well as OS module. Both of them work to collect static files from the Django applications into a single location that can easily be served in production of the project. The InSet lexicon dictionary was considered a static file as it has the Excel extension file. The static can be added to this project by importing the openpyxl module. The module helped to read and write Excel file and that will be saved into the workbook. Workbook is the container for all other parts of the document; it can manage the active sheet inside of the file. The author has separated the negative and positive words dictionary inside of two different sheets in one file, so openpyxl was needed to recognize those two sheets. The implementation of openpyxl can be seen in Figure 4.19.

```

# load InSet dictionary file
url =
os.path.join(os.path.dirname(os.path.dirname(__file__)), 'sentiment
/inset.xlsx')
inSetLexicon = load_workbook(url)
negatif = inSetLexicon['negatif']
positif = inSetLexicon['positif']
# tokenization
from nltk.tokenize import sent_tokenize, word_tokenize

# stemming
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
factory = StemmerFactory()
stemmer = factory.create_stemmer()

#stopword removal
from Sastrawi.StopWordRemover.StopWordRemoverFactory import
StopWordRemoverFactory
factory = StopWordRemoverFactory()
stopword = factory.create_stop_word_remover()

```

Figure 4.19 File and Library Prepared for Preprocessing

The code in Figure 4.20 shows the continuation of the previous code but this one is specified on the modules needed for the sentiment analysis. For the tokenization step, the module needed was the `nltk.tokenize`. It extracted the tokens from string of characters and returns the syllables from a single word. To do the stemming step, the author imported the Sastrawi library to help reducing inflected words in Bahasa Indonesia to the stem or base form. This library is widely used in many text analytics researches in Bahasa Indonesia. The next steps were punctuation removal and case folding. The function for those steps is shown in the figure below.

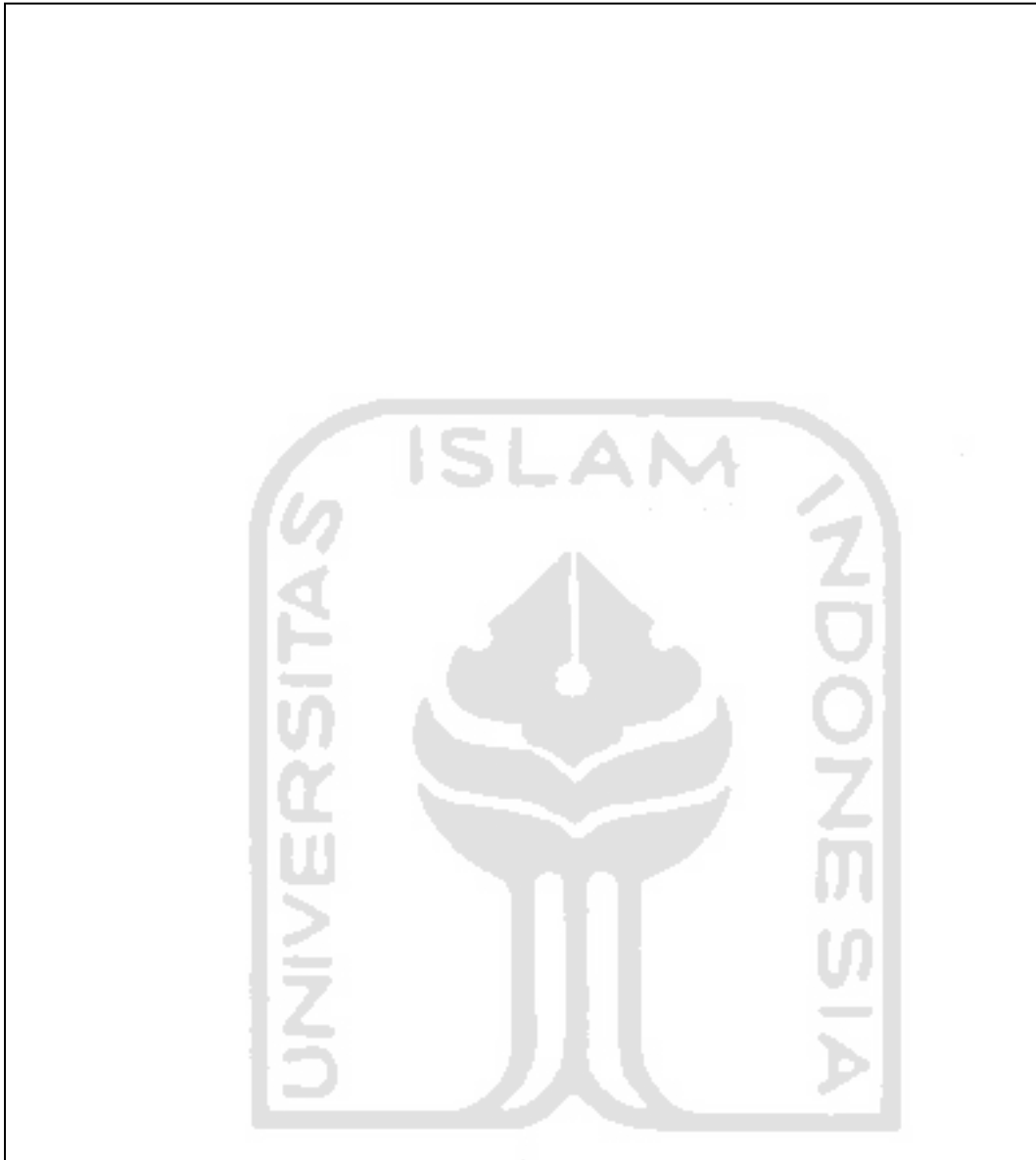


Figure 4.20 Function of Text Preprocessing

The following

Table 4.6 shows the example of data that will be used in the text preprocessing phase. The data below were obtained from research answer data table.

Table 4.6 Example of Sentence before Preprocessing

Mahasiswa dapat dengan mudah merasakan bahwa universitas memang telah

memiliki manajemen yang baik.
Mungkin kedepannya dapat diusahakan dapat memberi respon yang lebih cepat lagi.
Sistem ini memudahkan mahasiswa dalam kepengurusan surat resmi universitas.

After that, the sentence will be preprocessed using the functions mentioned above. The result of text preprocessing can be seen in Table 4.7.

Table 4.7 Preprocessed Sentence

'mahasiswa', 'mudah', 'rasa', 'universitas', 'memang', 'milik', 'manajemen', 'baik'
'mungkin', 'depan', 'usaha', 'beri', 'respon', 'lebih', 'cepat'
'sistem', 'mudah', 'mahasiswa', 'urus', 'surat', 'resmi', 'universitas'

4.4.2 Sentiment Analysis

In this phase, the system worked to look for the same words that are existing in both the data stored in the lexicon dictionary and preprocessed data. Then, for each sentiment matched words were given a score based on the weighting score in the lexicon. Each preprocessed word should have its own polarity score. The example of the result from assigning polarity score is shown in Table 4.8.

Table 4.8 Assign Polarity Score Result

['mudah', 4(-1)], ['memang', 4(-2)], ['milik', 1], ['baik', 3(-1)], ['rasa', (-2)]
['usaha', 1(-4)], ['beri', 2(-2)], ['respon', 2], ['lebih', 1], ['cepat', 3(-3)], ['mungkin', (-1)]
['mudah', 4(-1)], ['resmi', 2], ['sistem', (-4)]

After assigning polarity score, the last step to do was to calculate and count the total polarity score for both negative and positive sentiment classification. The calculation function of the total polarity score can be seen in Figure 4.21.

```
def sentimentScore(resultPositive, resultNegative):
    countPositive = 0
    countNegative = 0

    for arr in resultPositive:
        countPositive = countPositive + arr[1]
    for arr in resultNegative:
        countNegative = countNegative + arr[1]

    # print('countPositive: ', countPositive)
    # print('countNegative: ', countNegative)

    sentimentScore = countNegative + countPositive

    return countPositive, countNegative, sentimentScore
```

Figure 4.21 Function of Sentiment Scoring

The result after the calculation of sentiment score can be seen in

Table 4.9 and formula (4.1) down below to calculate the polarity score is as follows:

$$Sentiment\ Score = \sum_{i=1}^n overallPolarityScore(i) \quad (4.1)$$

Table 4.9 Sentiment Score

$(4-1) + (4-2) + 1 + (3-1) - 2 = 6$

$(1-4)+(2-2)+2+1+(3-3) -1 = -1$
$(4-1)+2-4 = 1$

The sentiment score that has been obtained from the calculation of polarity score was classified into three sentiment classification which are positive, negative, and neutral. In this system, the author implemented a function as a condition to classified the sentiment. If the sentiment score of a sentence or document > 0 then the sentence has a positive sentiment, if the sentiment score < 0 , then negative. Meanwhile, if the sentiment score $= 0$, then the sentence or document has a neutral sentiment. The function of sentiment classification can be seen in Figure 4.22 Function of Sentiment Classification.

```
def checkSentiment(scoreSentiment):
    result = ''
    if (scoreSentiment > 0):
        result = 'POSITIVE'
    elif (scoreSentiment < 0):
        result = 'NEGATIVE'
    elif (scoreSentiment == 0):
        result = 'NEUTRAL'

    return result
```

Figure 4.22 Function of Sentiment Classification

4.5 Testing

The testing method used to test the system being developed for this research is Black-box testing. This type of testing is also known as Behavioral testing that examines and tests the functionality of an application system without having to peer into its internal structures. Black-box testing can be applied virtually to every level of software testing like unit, integration, system and acceptance.

In this research, the testing phase were categorized into three different applications of the system. The landing, repository, and sentiment app has their own testing phase. The testing result by using Black-box method can be seen in Table 4.10

4.5.1 Landing App Testing

Table 4.10 Landing App Testing Result

Code	Test Case	Expected Result	Test Result
TC1-01	User fills in the sign up form to create new user account	System recorded the data and redirected to the sign in page	Success
TC1-02	User fills in the email that has been registered to the system in the sign up form	System showed a message below the email text field informing that the email has been registered	Success
TC1-03	User does not fill in one or more data in the required text fields	System showed a message to fill out the required text field	Success
TC1-04	User fills in the sign in form to log into the system	System authentication success and redirected to the workspace	Success
TC1-05	User fills in the invalid email address or wrong password	System showed a message and redirected to the sign in page	Success
TC1-06	User fills in the registration	System recorded the data of the user	Success

	form to register new workspace	as the registered and redirected to the workspace	
TC1-07	User fills in the join request form to join existing workspace	System recorded the data of the user with role and redirected to the workspace	Success
TC1-08	User fills in the wrong company workspace code	System showed a message and redirected to the join request page	Success

4.5.2 Repository App Testing

Table 4.11 Repository App Testing Result

Code	Test Case	Expected Result	Test Result
TC2-01	User accesses the workspace dashboard page	System showed the data in a number and chart displayed for the dashboard	Success
TC2-02	User fills in the research project form and submit to the system	System recorded the data and redirected to the management page	Success
TC2-03	User fills in the invalid email of person in charge or project member and submit to the system	System showed a message informing that email no found	Success
TC2-04	User fills in the email of project member twice (double) and submit to the system	System showed a message informing that email has been added	Success
TC2-05	User clicks the button delete to delete the research project data	System deleted the data and redirected to the management page	Success

Code	Test Case	Expected Result	Test Result
TC2-06	User clicks the button edit to edit the research project data	System redirected to the input form page with initial data to edit	Success
TC2-07	User fills in the research respondent form and submit to the system	System recorded the data and redirected to the management page	Success
TC2-08	User fills in the same name of research respondent in the form and submit to the system	System showed a message informing that name has been recorded before	Success
TC2-09	User does not fill in one or more data in the required text fields and submit to the system	System showed a message to fill out the required text field	Success
TC2-10	User clicks the button delete to delete the research respondent data	System deleted the data and redirected to the management page	Success
TC2-11	User clicks the button edit to edit the research respondent data	System redirected to the input form page with initial data to edit	Success
TC2-12	User fills in the research question form to the project and submit to the system	System recorded the research question and redirected to the research details page	Success
TC2-13	User clicks the button delete to delete the research question data	System deleted the data and redirected to the research details page	Success
TC2-14	User fills in the project member form and submit to the system	System recorded the data and redirected to the research details page	Success

Code	Test Case	Expected Result	Test Result
TC2-15	User fills in the invalid email of project member form and submit to the system	System showed a pop-up message informing that the email not found	Success
TC2-16	User clicks the button delete to delete the project member data	System deleted the data and redirected to the research details page	Success
TC2-17	User fills in research answer for each question to the answer form	System recorded the data and redirected to the research answer page	Success

4.5.3 Sentiment App Testing

Table 4.12 Sentiment App Testing Result

Code	Test Case	Expected Result	Test Result
TC3-01	User clicks the generate button to see sentiment analysis result	System showed the data in a number and chart displayed for the sentiment analysis	Success

CHAPTER 5. CONCLUSION AND RECOMMENDATION

5.1 Conclusion

In this thesis work, the SentiResearch web-based application has been developed with the implementation of sentiment analysis lexicon dictionary-based method. Based on what has been discussed earlier, there are several conclusions that can be attained. The conclusions are as follows:

1. The development of SentiResearch used Python programming language and Django framework. Within the development process, the Software Development Life Cycle (SDLC) process has been implemented to design, develop and test high quality application. Specifically, with Prototype method, it helped the users get a better understanding of the system being developed and helped the developer and/or designer to develop better application based on user's acceptance.
2. Sentiment Analysis Lexicon Dictionary-based method can be a great choice to implement with limited dataset resources. In this case, the dataset used to analyze the sentiment was the interview narrative or research answer data. Despite of having limited dataset, lexicon dictionary-based was able to calculate the sentiment score by providing a dictionary of words with weighting values. The dictionary used in the system is InSet lexicon as the only lexicon resources for sentiment analysis in Bahasa Indonesia (Indonesian official language).
3. The SentiResearch web-based application is dedicated to help UX researcher store the research data and get sentiment analysis insight from UX research project, especially a qualitative with in-depth interview research method.

5.2 Recommendation

Based on the development process and testing result, it is proved that this sentiment analysis web-based application has some limitations and shortcomings. Therefore, this application requires further development to add more features in order to develop a desirable application system. There are a number of suggestions and recommendations that can be implemented for future research and development of the system, such as:

1. Implement API to send email address validation to help improve the security of this application.
2. Add repository features that enables the user to store research media files like images, videos, and voice recordings.
3. Apply voice recognition program to identify words and phrases in spoken language from the interview activity and convert the audio to a machine-readable format like text.
4. Connect the system with Google calendar API for better scheduling experience.
5. In this research, the system was only tested with the black box testing method which aim to test the functional of the system. While the system also implemented sentiment analysis, there should be other testing method to test the calculation of sentiment score by using method like confusion matrix.

APPENDIX A. SURVEYS

ONLINE SURVEY

Online survey has been made by the author as one of the data collection methods. The survey has been designed to get insight from the potential user of the system, which is the research team in a company; that consists of the role of researcher, project manager, and the other roles of employees that can get access of the system.

The survey is divided into three sections, which are Roles and Responsibilities of The Researcher, The Needs of The Researcher, and The Demographic of Research Participant. The purpose in doing this online survey is to see the roles and needs of researchers in a qualitative research project, especially with in-depth interview method research.

Section 1 – The Roles and Responsibilities of The Researcher

The first section consists of questions that try to validate whether the qualitative research project is crucial to any research team. The results from each of the question show that most of the researcher has ever worked on a qualitative research project.

1. What kind of research you mostly do?

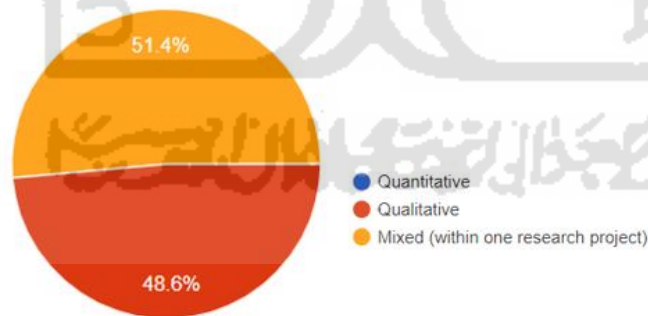


Figure A.1 Type of research Researcher mostly do

2. Which of these following research methods have you ever done?

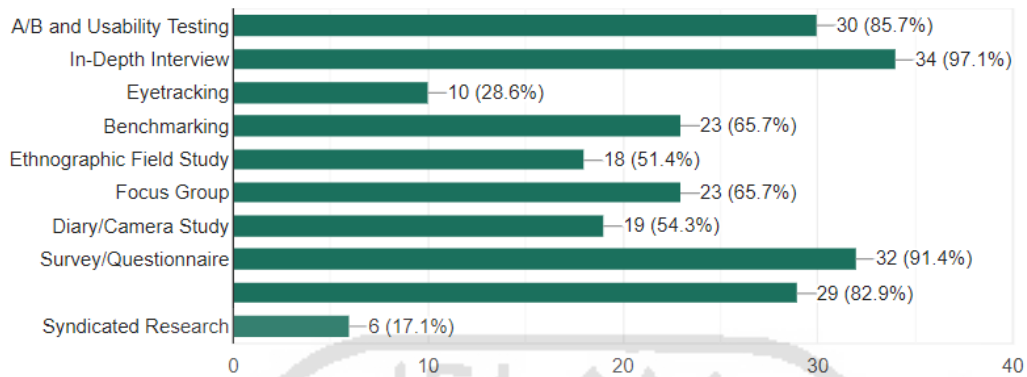


Figure A.2 Research method that have ever been done

3. Whom do you mostly get research requests from in your company?



Figure A.3 Role that would likely to request to do research

4. How often would you likely be assigned to do a qualitative research project? (specifically IDI method)

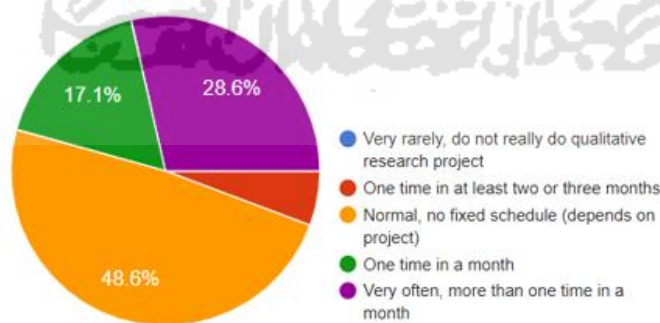


Figure A.4 Frequency of Qualitative Research

5. How many researchers does it take to complete one research project? (specifically IDI method)

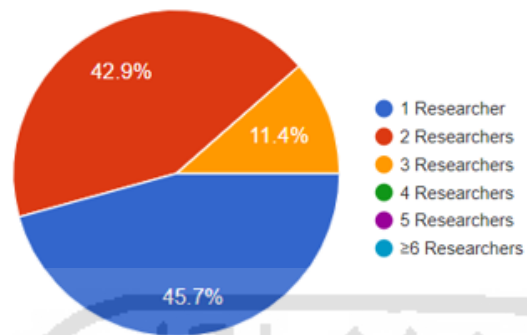


Figure A.5 Researcher needed in one research project

Section 2 – The Needs of The Researcher

The second section consists of questions that try to understand the needs of researcher in term of software and tool used while working on qualitative research project. The author has done a competitive analysis on the software and tool commonly used by the research team and tried to implement certain features in the development of the system.

1. Does your research team have specific research repository application system/platform to store research files?

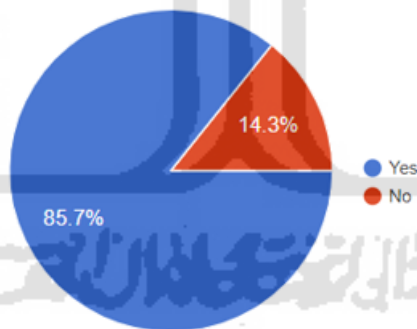


Figure A.6 Availability of research repository

2. Which of the following kinds of platform do you use to list down your research questions when designing Qualitative Research?

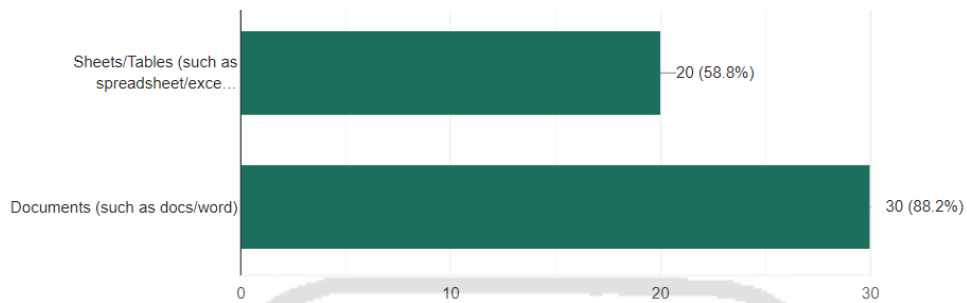


Figure A.7 Type of platform to design Qualitative Research

3. When working on an in-depth interview, which of these ways you prefer to take note or record?

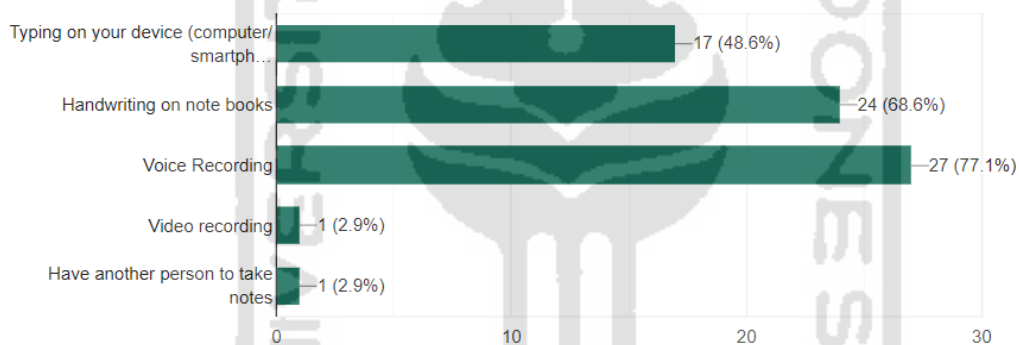


Figure A.8 Ways to take note or record an interview

4. What is the most common challenge do you face from research synthesis?

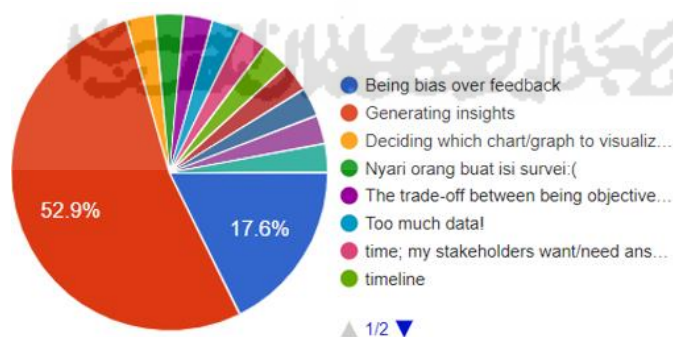


Figure A.9 Challenge in research synthesis

5. Please rate the importance of tools essential for your research team (Not Important – Very Important)

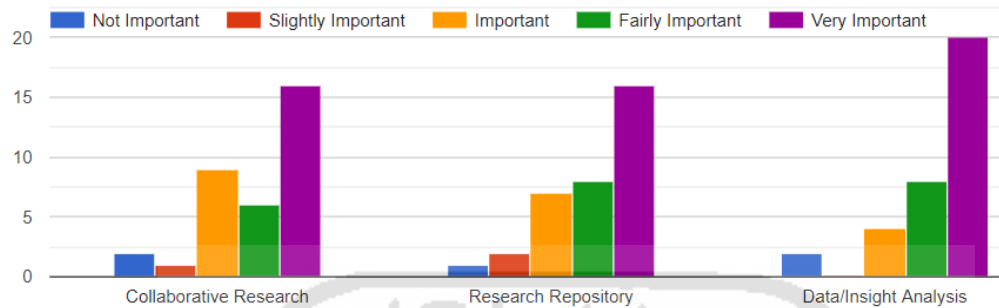


Figure A.10 The importance of tools

6. Please rate the importance of these features to be implemented in the application to support your research team (Not Important – Very Important)

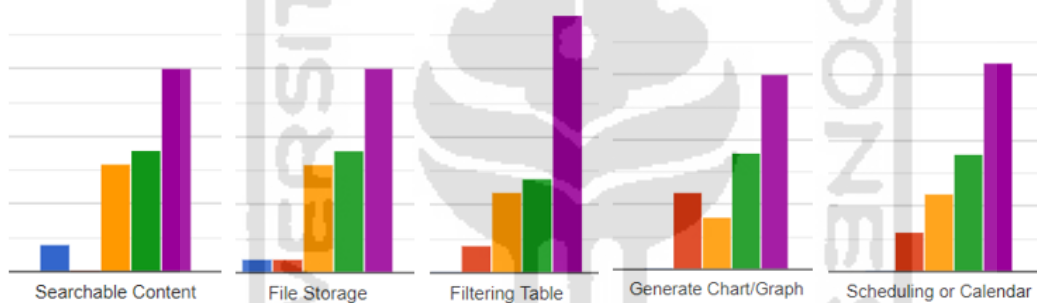


Figure A.11 The importance of features

Section 3 – The Demographic of Research Participant

The third and last section consists of questions that enable the author to gather some background information about potential users. Characteristics such as age, gender, and other information regarding of the company they are working at, and so on, are some of the examples of demographics used in this survey. The purpose of knowing the information about the company is to see how big and mature the company is to realize the need of a research team for the development.

1. What is your gender?

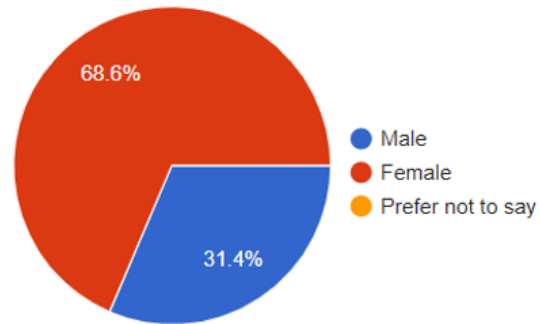


Figure A.12 Respondents' Gender

2. How old are you?

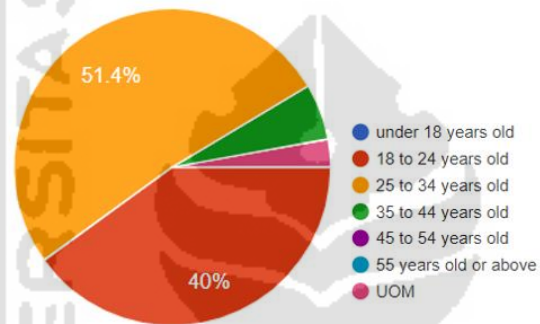


Figure A.13 Respondents' Age

3. How long have you been working as a researcher?

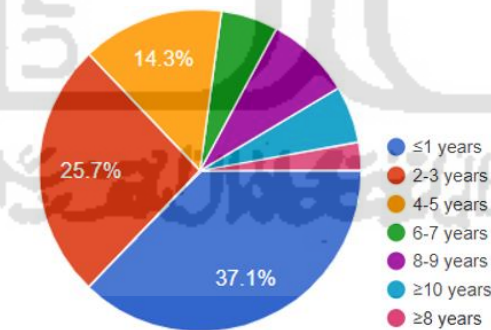


Figure A.14 Working period as a Researcher

4. How many employees does the company (you are working at) have?

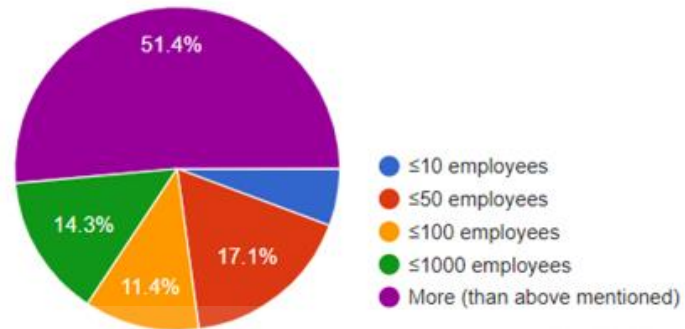


Figure A.15 Number of employees in the company

5. How many researchers does the company have?

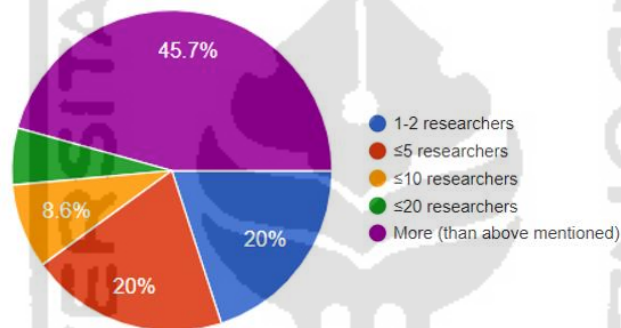


Figure A.16 Number of Researchers in the company

--- the end of survey ---

LIST OF REFERENCES

- Aman, S., & Szpakowicz, S. (2007). Identifying expressions of emotion in text. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4629 LNAI, 196–205.
https://doi.org/10.1007/978-3-540-74628-7_27
- Andreevskaia, A., Bergler, S., & Urseanu, M. (2007). All blogs are not made equal: Exploring genre differences in sentiment tagging of blogs. *ICWSM 2007 - International Conference on Weblogs and Social Media*.
- Aroomogan, K. (2015). *How Quant Traders Use Sentiment To Get An Edge On The Market*. Forbes.
<https://www.forbes.com/sites/kumesharoomoogan/2015/08/06/how-quant-traders-use-sentiment-to-get-an-edge-on-the-market/#4a9ddad44b5d>
- Asghar, M. Z., Kundi, F. M., Khan, A., & Ahmad, S. (2014). Lexicon-Based Sentiment Analysis in the Social Web. *J. Basic. Appl. Sci. Res*, 4(6), 238–248.
- Babich, N. (2017). *Prototyping 101: The Difference between Low-Fidelity and High-Fidelity Prototypes and When to Use Each* | Adobe Blog. Adobe Blog.
<https://theblog.adobe.com/prototyping-difference-low-fidelity-high-fidelity-prototypes-use/>
- Baccianella, S., Esuli, A., & Sebastiani, F. (2010). SENTIWORDNET 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. *Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010, November, 2200–2204*.
- Bollen, J., Mao, H., & Pepe, A. (2011). Modeling Public Mood and Emotion: Twitter Sentiment and Socio-Economic Phenomena. *Fifth International AAAI Conference on Weblogs and Social Media Modeling*, 10(5), 450–453.
- Gurusamy, V., & Kannan, S. (2016). *Preprocessing Techniques for Text Mining Preprocessing Techniques for Text Mining*. 5(October 2014), 7–16.
- Hancock, B. (2006). An Introduction to Qualitative Research Authors. *Qualitative Research*,

- 4th, 504. <https://doi.org/10.1109/TVCG.2007.70541>
- Holovaty, A., & Kaplan-Moss, J. (2009). *The Definitive Guide to Django: Web Development Done Right, Second Edition The Definitive Guide to Web Development Done Right Django 1.1 Django 1.1*. Apress.
- Indra, & Hartati, S. (2014). Aplikasi Pengolah Bahasa Alami untuk Info Gempa Bumi Terkini dengan Sumber Data pada Twitter @InfoBMKG. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, F-7-F-14. jurnal.uui.ac.id/index.php/Snati/article/download/3279/2952
- Jain, H. (2017). A Web Based Application for Sentiment Analysis. *International Journal of Education and Management Engineering*, 7(1), 25–35. <https://doi.org/10.5815/ijeme.2017.01.03>
- Jurek, A., Mulvenna, M. D., & Bi, Y. (2015). Improved lexicon-based sentiment analysis for social media analytics. *Security Informatics*, 4(1). <https://doi.org/10.1186/s13388-015-0024-x>
- Kadhim, A. I. (2018). An Evaluation of Preprocessing Techniques for Text Classification. *International Journal of Computer Science and Information Security*, 16(6), 22–32.
- Liu, B., Hu, M., & Cheng, J. (2005). Opinion Observer: Analyzing and Comparing Opinions on the Web. *Proceedings of the 14th International Conference on World Wide Web*, 342–351. <https://doi.org/10.1145/1060745.1060797>
- Maeda, J. (2015). *Design in Tech Report*. Kleiner Perkins Caufield & Byers. <https://designintech.report/2015/03/15/design-in-tech-report-2015/>
- Moebius, J. (2017). *2017 UX AND USER RESEARCH INDUSTRY REPORT FINDS SPIKES IN UX BUDGETS, TESTING FREQUENCY AND COMPETITOR RESEARCH IN 2016 / UserTesting*. UserTesting. <https://www.usertesting.com/about-us/press/press-releases/2017-ux-and-user-research-industry-report-finds-spikes-in-ux-budgets-in-2016>
- Neumann, P. (2004). *Prototyping*. 1–13.
- Nielsen, F.°Arup. (2011). *sentiment analysis in microblogs*. <http://arxiv.org/abs/1103.2903>

- Norman, D., & Nielsen, J. (2016). *The Definition of User Experience (UX)*. Nielsen Norman. <https://www.nngroup.com/articles/definition-user-experience/>
- Sadia, A., Khan, F., & Bashir, F. (2018). *An Overview of Lexicon-Based Approach For Sentiment Analysis. IEEEC*, 1–6.
- Sharma, N., Pabreja, R., Yaqub, U., Atluri, V., Ae Chun, S., & Vaidya, J. (2018). Web-based application for sentiment analysis of live tweets. *ACM International Conference Proceeding Series, November*, 1–3. <https://doi.org/10.1145/3209281.3209402>
- Shyam, A., & Mukesh, N. (2020). A Django Based Educational Resource Sharing Website: Shreic. *Journal of Scientific Research*, 64(01), 138–152. <https://doi.org/10.37398/jsr.2020.640134>
- Srinath, K. R. (2017). Python – The Fastest Growing Programming Language. *International Reasearch Journal of Engineering and Technology (IRJET)*, 4(12), 354–357. <https://www.technotification.com/2018/06/python-fastest-growing-programming-language.html>
- Westling, A., Brynielsson, J., & Gustavi, T. (2014). Mining the web for sympathy: The pussy riot case. *Proceedings - 2014 IEEE Joint Intelligence and Security Informatics Conference, JISIC 2014*, 123–128. <https://doi.org/10.1109/JISIC.2014.27>