KLIK: Kajian Ilmiah Informatika dan Komputer
ISSN 2723-3898 (Media Online)
Vol 4, No 1, Agustus 2023, Hal 144-151
DOI 10.30865/klik.v4i1.1075
https://djournals.com/klik

Search and Comparison of *Isim Ma'rifat* with Remove Diacritic on the Qur'an and Hadith of Bukhari

Dhafin Putra Aldi, Eko Darwiyanto, Moch. Arif Bijaksana*

School of Computing, Informatics Study Program, Telkom University, Bandung, Indonesia
Email: [1]dhafinputraaldi@students.telkomuniversity.ac.id, [2]ekodarwiyanto@telkomuniversity.ac.id,
[3,*]arifbijaksana@telkomuniversity.ac.id
Correspondence Author Email: arifbijaksana@telkomuniversity.ac.id

Abstract−The concept of Quranpedia was inspired by the idea of Wikipedia, which is the most popular online encyclopedia in the world. One of the main focuses of Quranpedia is to present data on Isim ma'rifat (definite nouns) found in the Quran and the Hadith Kutubus Sittah (six major Hadith collections). Isim ma'rifat refers to specific nouns or things categorized as nouns whose meanings are definite. One of the distinctive features of Isim ma'rifat is that it starts with the prefix AL (ال). Apart from the Quran, the explanations of Isim ma'rifat can also be found in the sayings of Prophet Muhammad (Hadith), including those in Sahih Bukhari. To obtain these Isim ma'rifat from the Quran and Hadith Bukhari, a search system for nouns with the prefix AL (ال) was developed using the remove diacritics method. The Java-based Isim detection program successfully found 9852 words with the prefix AL (ال) in the Quran and 70563 words in Hadith Bukhari. After removing duplicates and performing validation, 1041 Isim ma'rifat were identified in the Quran, and 1764 Isim ma'rifat were found in Hadith Bukhari. Then, a comparison was conducted, revealing that 376 Isim ma'rifat (ال) are present in both the Quran and Hadith Bukhari, 665 are only found in the Quran, and 1388 are exclusive to Hadith Bukhari. Based on this data, the calculation to measure Hadith Bukhari's ability to explain the Isim ma'rifat found in the Quran resulted in (376/1764)x100% = 21%. Hence, Hadith Bukhari can only explain 21% of the Isim ma'rifat (ال) present in the Quran.

Keywords: Quranpedia; Quran; Hadith of Bukhari; Isim ma'rifat; Remove Diacritics

1. INTRODUCTION

Wikipedia is the most popular online based encyclopedia in the world. Millions of people access Wikipedia daily to search for various information [1]. Building on the concept of Wikipedia, the idea of Quranpedia emerged—a website containing explanations about *Isim ma'rifat* (specific nouns) found in the Quran and the Hadiths of Kutubus Sittah. *Isim ma'rifat* refers to everything that can be categorized as tangible objects whose meanings are definite and can be perceived by everyone [2]. One characteristic of *Isim ma'rifat* is having the prefix AL (ال) [3]. The main source for explaining *Isim ma'rifat* is the Hadiths of the Prophet, including those from Hadith Bukhari. Therefore, a system is needed to search for *Isim ma'rifat* in the Quran and Hadith Bukhari. Similar findings of *Isim ma'rifat* in the Quran need to be compared with explanations in Hadith Bukhari to determine the extent of Hadith Bukhari's ability to explain *Isim ma'rifat* found in the Quran.

The previous research [1] discusses why people read Wikipedia. They successfully conducted a survey of 30,000 respondents and found patterns of behavior and motivations for people reading Wikipedia. The interesting thing about this research is the absence of topics related to the Quran and Hadiths based on the conducted survey. The research by [4] states that all words starting with "ال" are considered *Isim* or nouns, except for the letters of Muqata'at. According to [5], there are fourteen letters of *muqataat* (*Alif Lam Mim*, *Alif Lam Raa*, *Ta Ha*, *Ya*, *Sin*, etc). However, this needs to be validated regarding words with the prefix "ال" to determine whether there are words with the "ال" prefix that are not *Isim*. Furthermore, there is also research conducted by [3] aiming to determine the types of nouns in the Arabic language using the light stemming algorithm. The research successfully identified the types of words in the Arabic language, including *Isim* (nouns) and *Fi'il* (verbs), with an average accuracy rate of 97.54%. Unfortunately, this research did not include findings of *Isim* or *Fi'il* that could be used as a reference.

Based on the issue mentioned above, the author aims to propose a solution in the form of designing a system to search for *Isim ma'rifat* in the Quran and Hadith Bukhari with the prefix AL (ال). The Remove Diacritics method is used to facilitate the search for *Isim ma'rifat*, which involves removing diacritics or punctuation marks in the Arabic language. This algorithm will find all words with the prefix AL (ال) in the Quran and Hadith Bukhari verses. The search results will be displayed as data or information on the Quranpedia website.

2. RESEARCH METHODOLOGY

2.1 Research Steps

In this research, there are six stages to obtain the desired results. The outcome of this research is the data of *Isim ma'rifat* with prefix (ال), which will be used for the development of the Quranpedia website. The stages, such as algorithm or program design, validation process, evaluation, and others, can be seen in Figure 1 below.

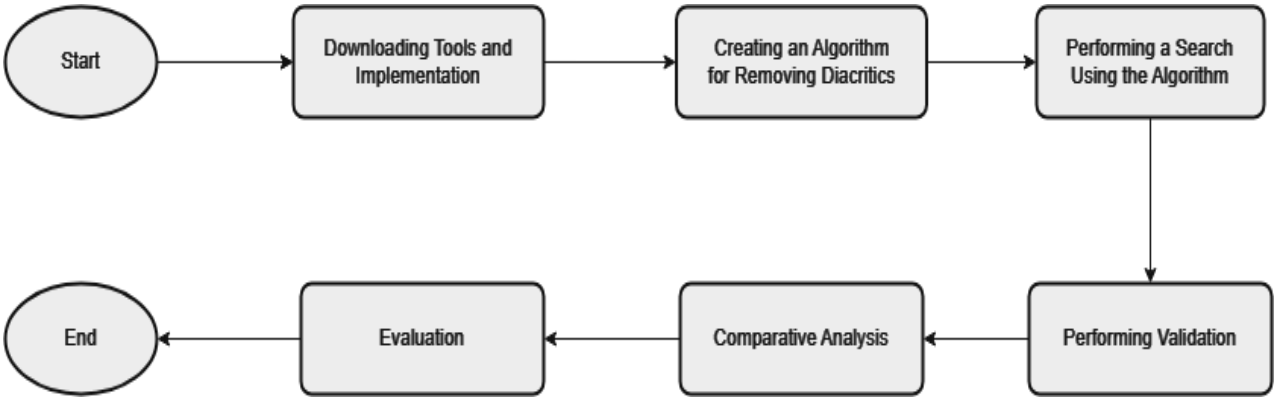


Figure 1. Research Flow

2.2 Downloading Tools and Implementation

This section explains all the devices and tools used in this research, starting from the device specifications, IDE, the database used, and others. More detailed explanations can be found in the table below. Table 1 explains the device specification, while Table 2 explains the tools used in this research.

Table 1. Device Specification

Name	Specification
Device	Razer Blade 15" Advanced (2019)
Processor	Intel® Core™ i7-8750H processor, 6 Cores / 12 Threads, 2.2GHz
Memory	16GB RAM DDR4-2667MHz
Operating System	Windows 11 Pro 64-bit Version 10.0.22621 Build 22621

Table 2. Tools and Implementation

Tools	Version	Download Source
Apache Netbeans	17	https://netbeans.apache.org/download/nb17
XAMPP	8.2.0	https://www.apachefriends.org/download.html
Java SE 8 (jdk 8)	8.0.3710.11	https://www.oracle.com/id/java/technologies/javase/javase8-archive-downloads.html
MySQL Connector	8.0.32	https://dev.mysql.com/downloads/connector/j/
Quran Database (public)	-	https://github.com/sinoridha/quran-indonesia-db
Shahih Bukhari Database (public)	-	https://github.com/irsyadulibad/hadits-database
Microsoft Excel	2302	https://www.microsoft.com/en-us/microsoft-365/download-office
Google Chrome	114.0.5735.199	https://www.google.com/chrome/

2.3 Creating an Algorithm for Removing Diacritics

Several steps are taken to create a necessary program for this research, such as creating a project, creating classes, using specific methods, and others. These steps can be seen in the Figure 2 below.

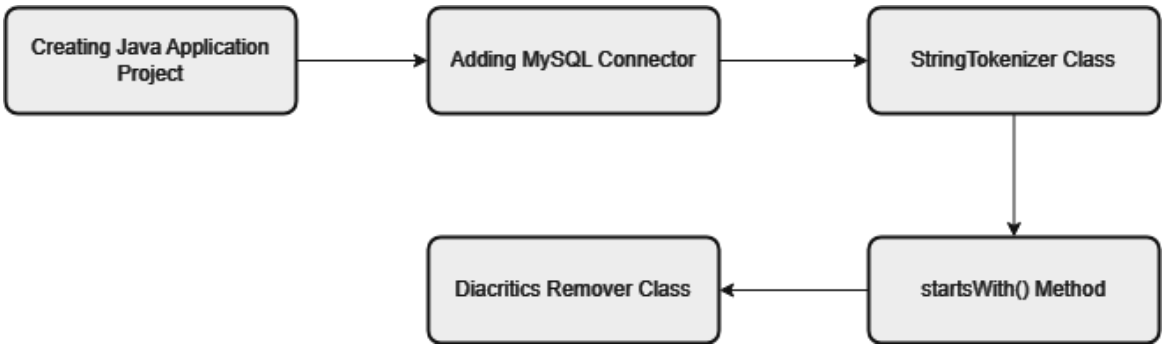


Figure 2. Creating main function steps

2.3.1 Creating Java Application Project

This stage is the first step in creating the program. A new project is created in the NetBeans IDE, and the subsequent steps will be added to this project.

2.3.2 Adding MySQL Connector

The MySQL Connector Java connects a MySQL database with a Java program. This is essential to establish a connection between the *Isim* search program and the Quran and Hadiths of Bukhari databases, allowing the program to search for and retrieve *Isim* data from the Quran and Hadiths of Bukhari databases.

2.3.3 StringTokenizer Class

In text mining, tokenizing is a process that converts a sentence into several words (strings) [6]. In Java, the String Tokenizer class is useful for breaking a string into several tokens. This class will be used in the program's development to tokenize verses from the Quran and Hadiths of Bukhari obtained from the database.

2.3.4 startsWith() Method

This method is a string class method used to check the prefix of a string. This method is necessary to find words with the prefix (ال). After using the StringTokenizer class and applying this method, the program will tokenize all verses that have the prefix (ال).

2.3.5 Diacritics Remover Class

The study of Quranic sciences explains that the knowledge that deals with diacritics in the Quran is called "*al-dabt*" or "*naqt wa syakal*," which is closely related to the "*rasm mushaf*" or the study of the Quranic script. This is because orthography between letters and diacritics facilitates someone reading the Quran [7]. This class functions to remove punctuation or diacritics. The purpose of removing diacritics is to facilitate the process of searching for *Isim ma'rifat* (ال) and minimize duplication. For example, in the word اللهِ (Allah), if the diacritics are not removed, the program may find اللهَ (Allaha), اللهِ (Allahi), and اللهُ (Allahu), which all have the same meaning, Allah. This can add unnecessary duplication in the findings of *Isim ma'rifat* (ال).

2.4 Performing a Search Using the Algorithm

After creating the search program, it will be executed. The output generated from the program will then be processed and validated. All the output results from the Quran and Hadiths of Bukhari will be copied into Microsoft Excel for various purposes, such as counting the number of occurrences, removing duplicates, and other operations as needed.

2.5 Performing Validation

According to a previous study [4], all words that start with (ال) are considered *Isim*. Therefore, this needs to be validated by conducting validation on the program's search results. The Quranic findings are validated manually using the website quran.bblm.go.id, and for Hadiths of Bukhari, context.reverso.net is used. Quran BBLM is a website that provides information about the vocabulary explanations in the verses of the Quran [8]. Reverso Context is a language portal launched in France that provides online translation aids and language services, namely online dictionaries, online bilingual concordances, spell checking, and conjugation tools and is used by 96 million monthly active users [9]. The validation results will be copied into Microsoft Excel for further processing.

2.6 Comparative Analysis

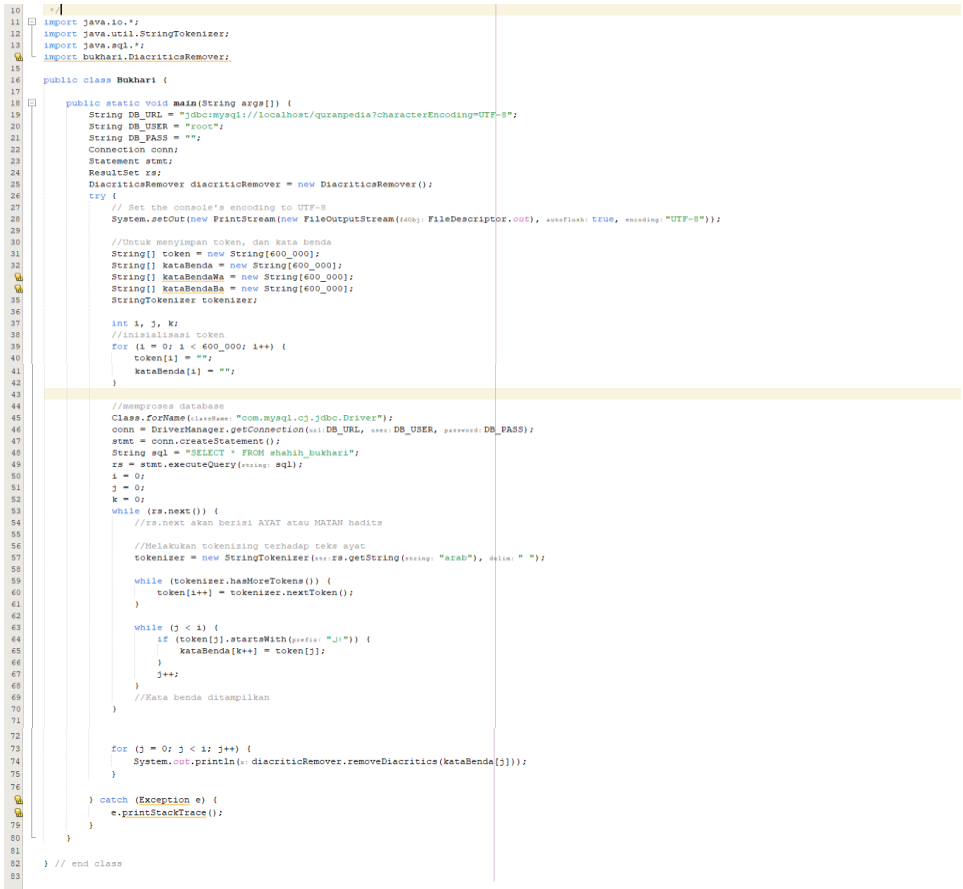
This stage is carried out after the validation process is completed. Valid *Isim ma'rifat* (ال) data from the Quran and Hadiths of Bukhari will be transferred to Microsoft Excel for further analysis. This process aims to find similarities and differences between the *Isim ma'rifat* (ال) found in the Quran and the *Isim ma'rifat* (ال) found in the Hadiths of Bukhari. It also aims to determine how Hadiths of Bukhari can effectively explain the *Isim ma'rifat* (ال) in the Quran.

2.7 Evaluation

In this stage, an analysis is conducted on all the findings, starting from the data of the program's entire output, data after removing duplicates, data after validation, and data after comparison. This process is also crucial to measure the accuracy of the algorithm or program that has been created. The *Isim ma'rifat* (ال) data will be implemented in developing the Quranpedia website.

3. RESULT AND DISCUSSION

3.1 Remove Diacritics Algorithm



```
*/
import java.io.*;
import java.util.StringTokenizer;
import java.sql.*;
import bukhari.DiacriticsRemover;

public class Bukhari {

    public static void main(String args[]) {
        String DB_URL = "jdbc:mysql://localhost/quranpedia?characterEncoding=UTF-8";
        String DB_USER = "root";
        String DB_PASS = "";
        Connection conn;
        Statement stmt;
        ResultSet rs;
        DiacriticsRemover diacriticRemover = new DiacriticsRemover();
        try {
            // Set the console's encoding to UTF-8
            System.setOut(new PrintStream(new FileOutputStream(fdObj: FileDescriptor.out), autoFlush: true, encoding: "UTF-8"));

            //Untuk menyimpan token, dan kata benda
            String[] token = new String[600_000];
            String[] kataBenda = new String[600_000];
            String[] kataBendaWa = new String[600_000];
            String[] kataBendaBa = new String[600_000];
            StringTokenizer tokenizer;

            int i, j, k;
            //inisialisasi token
            for (i = 0; i < 600_000; i++) {
                token[i] = "";
                kataBenda[i] = "";
            }

            //memproses database
            Class.forName(className: "com.mysql.cj.jdbc.Driver");
            conn = DriverManager.getConnection(url:DB_URL, user:DB_USER, password:DB_PASS);
            stmt = conn.createStatement();
            String sql = "SELECT * FROM shahih_bukhari";
            rs = stmt.executeQuery(string: sql);
            i = 0;
            j = 0;
            k = 0;
            while (rs.next()) {
                //rs.next akan berisi AYAT atau MATAN hadits

                //Melakukan tokenizing terhadap teks ayat
                tokenizer = new StringTokenizer(str:rs.getString(string: "arab"), delim: " ");

                while (tokenizer.hasMoreTokens()) {
                    token[i++] = tokenizer.nextToken();
                }

                while (j < i) {
                    if (token[j].startsWith(prefix: "Ø§")) {
                        kataBenda[k++] = token[j];
                    }
                    j++;
                }
                //Kata benda ditampilkan
            }


            for (j = 0; j < i; j++) {
                System.out.println(x: diacriticRemover.removeDiacritics(kataBenda[j]));
            }

        } catch (Exception e) {
            e.printStackTrace();
        }
    }

} // end class
```

Figure 3. Main function

To obtain *Isim ma'rifat* with the prefix AL (ال), the method used is **startsWith()** as seen in Figure 3 (line 64), which checks whether the string value starts with the specified prefix, in this case using the prefix (ال). The program will tokenize each verse when it finds the prefix (ال). The search is performed in the Quran and Hadith Bukhari databases.

```
*/
import java.util.ArrayList;

public class DiacriticsRemover {

    // remove all diacritics from text
    public String removeDiacritics(String currentText) {
        ArrayList<String> tokens = new ArrayList<String>();
        StringBuffer modifiedWord = new StringBuffer(str:"");
        ArrayList<String> modifiedText = removeExtraSpaces(currentText);

        // for each token in the text
        for (int i = 0; i < modifiedText.size(); i++) {
            modifiedWord.setLength(newLength: 0);
            String ctoken = modifiedText.get(index: i);
            for (int j = 0; j < ctoken.length(); j++) {
                if (!(Constants.diacritics.contains(x: ctoken.substring(beginIndex: j, j + 1)))) {
                    modifiedWord.append(str:ctoken.substring(beginIndex: j, j + 1));
                }
            }
            tokens.add(e: modifiedWord.toString());
        }

        StringBuilder result = new StringBuilder();

        for (String t : tokens) {
            result.append(str:t).append(str:" ");
        }

        if (result.length() > 0) {
            result.setLength(result.length() - 1);
        }

        return result.toString();
    }

    private ArrayList<String> removeExtraSpaces(String currentText) {
        ArrayList<String> tt = new ArrayList<String>();
        StringBuilder word = new StringBuilder();
        currentText = currentText + " ";

        for (int i = 0; i < currentText.length(); i++) {
            // if the character is not a space, add it to a word
            if (!Character.isWhitespace(ch: currentText.charAt(index: i))) {
                word.append(c: currentText.charAt(index: i));
            } else {
                if (word.length() != 0) {
                    tt.add(e: word.toString());
                    word.setLength(newLength: 0);
                }
            }
        }

        return tt;
    }
}
```

Figure 4. Diacritics Remover

Figure 4 above is the Diacritics Remover Class. This class is used to remove diacritics or harakat from the program's output. The purpose of removing diacritics is to simplify the process of searching for *Isim ma'rifat* (ال) and minimize unnecessary duplication. For example, for the word الله (*Allah*), if the diacritics are not removed, the program will find اللهَ (*Allaha*), للهِ (*Allahi*), and اللهُ (*Allahu*), all of which have the same meaning, Allah. This could lead to unnecessary duplication in the findings of *Isim ma'rifat* (ال).

```
import java.util.ArrayList;

public class Constants {

    public static final char ALIF = '\u0627';
    public static final char ALIF_HAMZA_ABOVE = '\u0623';
    public static final char ALIF_HAMZA_BELOW = '\u0625';
    public static final char ALIF_MADDA = '\u0622';
    public static final char TAA_MARBUTA = '\u0629';
    public static final char HAA = '\u0647';
    public static final char YAA = '\u064A';
    public static final char DOTLESS_YAA = '\u0649';
    public static final String TATWEEL = "\u0640";

    public static ArrayList<String> punctuations = new ArrayList<String>();
    public static ArrayList<String> diacritics = new ArrayList<String>();

    static {
        punctuations.add(",");
        punctuations.add(":");
        punctuations.add("،");    // Arabic comma
        punctuations.add("؛");    // Arabic semi-colons
        punctuations.add("؟");    // Arabic question mark
        punctuations.add("'");
        punctuations.add("!");
        punctuations.add("?");
        punctuations.add("-");
        punctuations.add("--");
        punctuations.add("(");
        punctuations.add(")");
        punctuations.add("..."); // points of ellipsis.

        diacritics.add("\u064B");  // FATHATAN
        diacritics.add("\u064C");  // DAMMATAN
        diacritics.add("\u064D");   // KASRATAN
        diacritics.add("\u064E");  // FATHA
        diacritics.add("\u064F");  // DAMMA
        diacritics.add("\u0650");  // KASRA
        diacritics.add("\u0651");  // SHADDA
        diacritics.add("\u0652");  // SUKUN
    }
}
```

Figure 5. Constant

In this class, encoding is performed on each verse finding so that it no longer uses diacritics or punctuation when the verse appears. This is achieved by removing excessive spacing in the Arabic text and eliminating all harakat or diacritics that commonly appear, such as *FATHA*, *DAMMA*, *KASRA*, *SHADDA*, *SUKUN*, and others, as seen in Figure 5. To execute the remove diacritics method, the code above needs to be turned into a library named Constant.

3.2 Search Result

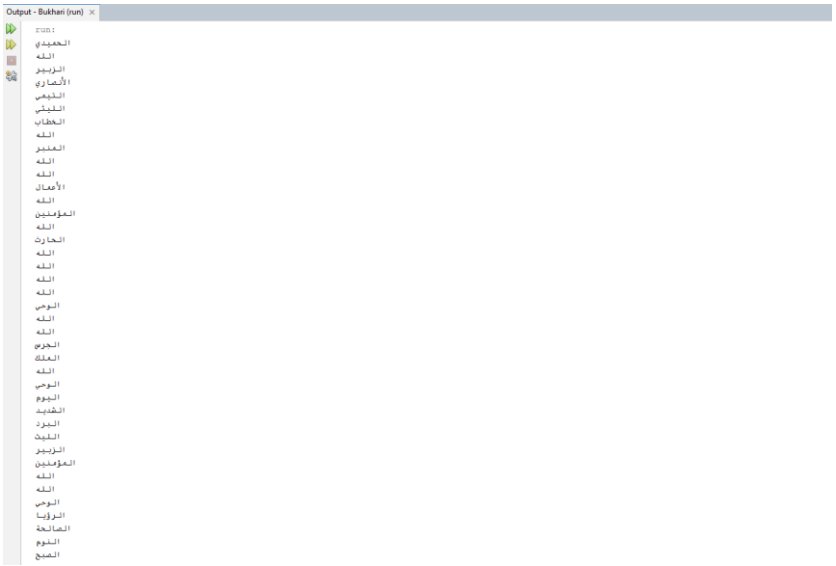


Figure 6. Program Output

Table 3. Sample of the Total Program Output

Prefix (ال)			
Number (Al-Quran)	Number (Hadith of Bukhari)	Al-Quran	Hadith of Bukhari
1	1	الرا	الابن
2	2	الرا	الابن
3	3	الرا	الابن
4	4	الرا	الابنة

5	5	الٓرٰ	الاثنتين
6	6	الٓمٓ	الاثنين
7	7	الٓمٓ	الاثنين
8	8	الٓمٓ	الاثنين
9	9	الٓمٓ	الاثنين
10	10	الٓمٓ	الاثنين
n - 9851	n - 70563	…	…

While Figure 6 shows the original program output, Table 3 above presents sorted data from the search for *Isim ma'rifat* (ال) using the remove diacritics algorithm in the Quran and Hadith Bukhari databases. The program found 9851 data in the Quran and 70563 data in Hadith Bukhari, which still need to be checked to ensure that they are valid *Isim ma'rifat* (ال) and that there are no duplicates.

Table 4. Sample of the Unique Program Output

Prefix (ال)			
Number (Al-Quran)	Number (Hadith of Bukhari)	Al-Quran	Hadith of Bukhari
1	1	الٓرٰ	الابن
2	2	الٓمٓ	الابنة
3	3	الٓمٓرٰ	الاثنتين
4	4	الٓمٓصٓ	الاثنين
5	5	الٓنٓ	الاجتناب
6	6	الٓيٓ	الاحتباء
7	7	الٓتي	الاحتلام
8	8	الأثمين	الاختلاف
9	9	الأخر	الاختلاق
10	10	الأخرين	الاستبضاع
n - 1183	n - 3306	…	…

Table 4 above shows the unique Quran and Hadith Bukhari data after removing duplicates. In the Quran, out of 9851 data, 1183 unique data were found. Meanwhile, in Hadith Bukhari, out of 70563 data, 3306 unique data were found. This indicates the presence of many duplicate data in the Quran and Hadith Bukhari. Most of the removal of duplicates was done automatically using Microsoft Excel. The following data still needs validation to determine their accuracy.

3.3 Validation Results

Table 5. Sample Results of Validation Using Quran BBLM and Context Reverso

Isim ma'rifat (ال)			
Number (Al-Quran)	Number (Hadith of Bukhari)	Al-Quran	Hadith of Bukhari
1	1	الأثمين	الابن
2	2	الأزفة	الابنة
3	3	الأفلين	الاثنين
4	4	الأمرون	الاجتناب
5	5	الأمنين	الاحتلام
6	6	الابتر	الاختلاف
7	7	الابرار	الاستبضاع
8	8	الابصار	الاستخارة
9	9	الابل	الاستسقاء
10	10	الابواب	الاستغفار
n - 1041	n - 1764	…	…

Table 5 above presents the validation results of *Isim ma'rifat* (ال). The validation process uses the quran.bblm.go.id and context.reverso.net websites. As a result, 1041 *Isim ma'rifat* (ال) were found in the Quran, and 1764 *Isim ma'rifat* (ال) were found in Hadith Bukhari. These data will be used to compare and find similarities between *Isim ma'rifat* (ال) in the Quran and Hadith Bukhari.

Table 6. Analysis of Validation Results

Validation Results of the Quran	
Data	1183
Isim (noun)	1041
Not a noun	142
Accuracy	88%

Validation Results of Hadith of Bukhari	
Data	3306
Isim (noun)	1764
Not a noun	1542
Accuracy	53%

Table 6 above represents the analysis of validation results for the *Isim ma'rifat* (ال) in the Quran and the Hadiths of Bukhari. In the Quran, out of 1183 data, 1041 valid *Isim ma'rifat* (ال) were found with a high accuracy rate of 88%. On the other hand, in the Hadiths of Bukhari, out of 3306 data, 1764 valid *Isim ma'rifat* (ال) were found with an accuracy rate of 53%.

3.4 Comparative Analysis Results

Table 7. Sample of Similarities and Differences of *Isim ma'rifat* (ال) in the Quran with Hadith Bukhari

Isim ma'rifat (ال)					
Number	Similarities between the Quran and Hadith of Bukhari	Number	Only found in Quran	Number	Only found in Hadith of Bukhari
1	الابصار	1	الأثمين	1	الابن
2	الابل	2	الأزفة	2	الابنة
3	الابواب	3	الأفلين	3	الاثنين
4	الاثم	4	الأمرون	4	الاجتناب
5	الاحاديث	5	الأمنين	5	الاحتلام
6	الاحزاب	6	الابتر	6	الاختلاف
7	الاحسان	7	الابرار	7	الاستبضاع
8	الاحلام	8	الابيض	8	الاستخارة
9	الاحمال	9	الاثيم	9	الاستسقاء
10	الاخ	10	الاجداث	10	الاستغفار
n - 376	…	n - 665	…	n - 1388	…

Table 7 above shows similar and different data in the Quran and Hadith Bukhari. After comparison, 376 *Isim ma'rifat* (ال) were found to be the same in the Quran and Hadith Bukhari. Additionally, 665 *Isim ma'rifat* (ال) in the Quran were not found in Hadith Bukhari, while 1388 *Isim ma'rifat* (ال) in Hadith Bukhari were not found in the Quran. This indicates Hadith Bukhari's ability to explain only 21% of the Quran's Arabic nouns or *Isim ma'rifat* (ال). Furthermore, 1388 *Isim ma'rifat* (ال) in Hadith Bukhari is not found in the Quran.

Table 8. Data Processing Results

Isim ma'rifat (ال)	Total
Isim found in the Quran	1041
Isim found in Hadith of Bukhari	1764
Isim found in both (similarities)	376
Isim found only in the Quran	665
Isim found only in Hadith of Bukhari	1388

Table 8 above shows the values of the data in this study. These values were obtained after eliminating duplicates, validation, and comparison. These data will be used to find the percentage of shared and unique data between the Quran and Hadiths.

Table 9. Quran Data Percentage

Percentage of Matching and Non-Matching Data in the Quran	Total
Number of matching data in the Quran	376
Number of non-matching data in the Quran	665
Percentage of matching data in the Quran	36%
Percentage of non-matching data in the Quran	64%

Table 9 concludes that in the Quran, 665 data were found to be different from the data in the Hadiths of Bukhari. The Quran has 36% of data that matches with the Hadiths and 64% of data that does not match. These 665 non-matching *Isim ma'rifat* (ال) words are not explained in the Hadiths of Bukhari.

Table 10. Hadith of Bukhari Data Percentage

Percentage of Matching and Non-Matching Data in Hadiths of Bukhari	Total
Number of matching data in Hadiths of Bukhari	376
Number of non-matching data in Hadiths of Bukhari	1388
Percentage of matching data in Hadiths of Bukhari	21%
Percentage of non-matching data in Hadiths of Bukhari	79%

On the other hand, Table 10 shows in the Hadiths of Bukhari, 1388 data were found to be different from the Quran. The percentage of data that matches the Quran in Hadiths of Bukhari is 21%, while the percentage of non-matching data

is 79%. This shows that Hadiths of Bukhari can only explain 21% of the Nouns or *Isim ma'rifat* (ال) found in the Quran. Additionally, 1388 *Isim ma'rifat* (ال) in the Hadiths of Bukhari are not found in the Quran.

4. CONCLUSION

Based on the calculations above, the conclusion is that the created program successfully finds *Isim ma'rifat* (ال) in the Quran and Hadith Bukhari with an accuracy level of 88% in the Quran and 53% in Hadith Bukhari. After conducting the comparison analysis, it can be concluded that Hadith Bukhari can only explain 21% or 376 *Isim ma'rifat* (ال) found in the Quran, while 665 *Isim ma'rifat* (ال) are not explained in Hadith Bukhari, and there are 1388 *Isim ma'rifat* (ال) in Hadith Bukhari that are not found in the Quran. Furthermore, the statement from the research [4] that all words starting with "ال" are *Isim* (except for 14 letters of *Muqata'at*) has been refuted, as based on the findings of this research, 142 words are starting with "ال" that are not considered *Isim* in the Quran, and 1542 words starting with "ال" that are not considered *Isim* in the Hadiths of Bukhari.

REFERENCES

[1] P. Singer et al., "Why we read Wikipedia," in 26th International World Wide Web Conference, WWW 2017, International World Wide Web Conferences Steering Committee, 2017, pp. 1591–1600. doi: 10.1145/3038912.3052716.

[2] A. A. Qifari, "Nakirah dan Ma' Rifah Fii Al-Qur'an," Jurnal Shaut Al-'Arabiyah, vol. 10, no. 1, pp. 107–114, 2022, doi: 10.24252/saa.v10i1.29432.

[3] N. Maliya and R. M. Pratama, "SISTEM DETEKSI JENIS KATA DALAM BAHASA ARAB DENGAN MENGGUNAKAN ALGORITMA LIGHT STEMMING," Universitas Muhammadiyah Jember, pp. 1–6, 2018.

[4] Q. A. Al-Radaideh and M. Afif, "Arabic Text Summarization Using Aggregate Similarity Arabic Natural Language Processing View project ARABIC TEXT SUMMARIZATION USING AGGREGATE SIMILARITY," in The 2009 International Arab Conference on Information Technology (ACIT'2009), 2009, pp. 1–8. doi: 10.13140/2.1.4173.1209.

[5] M. Saeed and A. Hayat, "Difference of Opinions among Scholars in the Explanation and Interpretation of Al- Huruful Muqataat: A Comparative Study," The Journal of Religious Studies of the University of Chitral (UOCHJRS), vol. 3, no. 2, 2020, doi: 10.33195/uochjrs-v3i21952020.

[6] A. Hassani, A. Iranmanesh, and N. Mansouri, "Text Mining using Nonnegative Matrix Factorization and Latent Semantic Analysis," Shahid Bahonar University of Kerman, vol. v3, pp. 1–26, Nov. 2020, [Online]. Available: http://arxiv.org/abs/1911.04705

[7] N. R. I. Hula and B. Arsyad, "DIAKRITIK AL-QUR'AN MENURUT PREFERENSI ABU DAWUD," 'A Jamiy: Jurnal Bahasa dan Sastra Arab, vol. 9, no. 2, pp. 265–284, 2020, doi: 10.31314/ajamiy.9.2.265-284.2020.

[8] A. M. Permana, E. Darwiyanto, and A. M. Bijaksana, "Pembobotan dan Pemeringkatan Ayat Al-Quran Berdasarkan Compound, Term Frequency dan Prinsip Pareto Untuk Membantu Hafalan," e-Proceeding of Engineering, vol. 8, no. 2, pp. 3352–3359, 2021.

[9] B. Zemni and C. Bounaas, "Audiovisual Translation and contextual dictionaries: An exploratory comparative study of Reverso Context and Almaany uses," Asian EFL Journal, vol. 27, no. 5, pp. 274–309, 2020, [Online]. Available: https://www.researchgate.net/publication/348150615

[10] V. Amrizal, R. Melita, B. H. Suseno, and T. Dirjam, "PENERAPAN METODE TERM FREQUENCY INVERSE DOCUMENT FREQUENCY (TF-IDF) DAN COSINE SIMILARITY PADA SISTEM TEMU KEMBALI INFORMASI UNTUK MENGETAHUI SYARAH HADITS BERBASIS WEB (STUDI KASUS: SYARAH UMDATIL AHKAM)," JURNAL TEKNIK INFORMATIKA, vol. 11, no. 2, pp. 149–164, Nov. 2018, doi: 10.15408/jti.v11i2.8623.

[11] S. Kumar, "A REVIEW ON CLIENT-SERVER BASED APPLICATIONS AND RESEARCH OPPORTUNITY," Article in International Journal of Scientific Research, 2019, doi: 10.24327/ijrsr.2019.1007.3768.

[12] A. N. Sa'adah, "COMPARATIVE STUDY BETWEEN ENGLISH AND ARABIC: COMMON NOUN VS ISIM NAKIRAH AND PROPER NOUN VS ISIM MA'RIFAH AND ITS IMPLICATION IN ENGLISH AND ARABIC LANGUAGE TEACHING," IAIN JEMBER, pp. 1–137, 2020.

[13] N. Zaman, "MAKNA HURUF MUQATHA'AH DI DALAM AL QURAN," INSTITUT PERGURUAN TINGGI ILMU AL-QUR'AN JAKARTA, pp. 1–73, 2022.

[14] N. Omar and Q. Al-Tashi, "Arabic nested noun compound extraction based on linguistic features and statistical measures," GEMA Online Journal of Language Studies, vol. 18, no. 2, pp. 93–107, May 2018, doi: 10.17576/gema-2018-1802-07.

[15] P. D. Hanifah, "MENUMBUHKAN SIKAP ILMIAH MELALUI KAJIAN TEMATIK SAINS QS. AL-MU'MINUN AYAT 12-14 MENGGUNAKAN TEKNIK QUR'AN JURNAL," Jurnal Kajian Pendidikan Sains, vol. 6, no. 2, 2020.

[16] K. Darwish, H. Mubarak, and A. Abdelali, "Arabic Diacritization: Stats, Rules, and Hacks," in Proceedings ofThe Third Arabic Natural Language Processing Workshop (WANLP), 2017, pp. 9–17.

[17] K. Shimonaka, Y. Higo, J. Matsumoto, K. Naito, and S. Kusumoto, "Towards automated generation of Java methods: A way of automated reuse-based programming," in 2018 IEEE 12th International Workshop on Software Clones, IWSC 2018 - Proceedings, Institute of Electrical and Electronics Engineers Inc., Mar. 2018, pp. 30–36. doi: 10.1109/IWSC.2018.8327316.

[18] B. A. Adanan, "KITAB SHAHIH AL-BUKHARI," Ihya Al-Arabiyah: Jurnal Pendidikan Bahasa dan Sastra Arab, vol. 6, no. 2, pp. 1–15, 2020.

[19] Z. M. Mubarok, Nurkholis, and Irham Muhammad, "Analisis Isim Ma'rifat dalam Al-Qur'an Surat Ash-Shaff," Al Mitsali: Jurnal Penelitian dan Pendidikan Bahasa Arab, vol. 1, no. 1, pp. 52–60, 2021.

[20] F. A. S. Jaya, "AL-QUR'AN DAN HADIS SEBAGAI SUMBER HUKUM ISLAM," JURNAL INDO-ISLAMIKA, vol. 9, no. 2, pp. 204–216, 2019.