Information Retrieval System in the Bible

Apriandy Angdresey[1], Miguel Angelo Lamongi[2], Rinaldi Munir[3]
[1,3]Department of Informatics Engineering, Universitas Katolik De La Salle, Manado
[2]School of Electrical Engineering and Informatics, Institut Teknologi Bandung
Email: **[1]aangdresey@unikadelasalle.ac.id**, [2]14013023@unikadelasalle.ac.id, [3]rinaldi@informatika.org

Abstrak

Sistem temu kembali informasi digunakan untuk mencari dokumen yang relevan sehingga dapat diperoleh dengan cepat dan tepat. Banyak orang Kristen yang ingin mempelajari dan mendalami Injil, namun seringkali mengalami kendala dalam mencari ayat-ayat dalam Injil dan topik yang berkaitan dengan kebutuhan yang dicari oleh pengguna. Hal ini dikarenakan, pengguna harus mencari satu per satu setiap ayat dalam keempat Injil. Dalam penelitian ini, penulis menggunakan ayat-ayat Alkitab dalam Injil sebagai dokumen, sehingga ayat-ayat tersebut dapat dicari berdasarkan tingkat relevansinya atau kemiripannya dengan kata kunci yang dimasukkan oleh pengguna. Selanjutnya untuk mengetahui tingkat relevansi antara dokumen dengan kata kunci, dihitung menggunakan Vector Space Model. Berdasarkan aplikasi yang telah berhasil dibangun, aplikasi ini dapat menampilkan 10 dokumen yang terkait dengan kata kunci yang dicari dan diurutkan dari yang paling relevan, dengan nilai kemiripan tertinggi yaitu 78,65%.

Kata kunci - Temu Kembali Informasi, Vector Space Model, Alkitab.

Abstract

Information retrieval is used to search for relevant documents so that they can be obtained quickly and precisely. There are many Christians who want to study the Gospel. However, often experience problems in finding the Gospel verse and topics dealing with the need to search by the user, due to having to search one by one each verse in the four Gospels. In this study, the authors used the Bible verses in the Gospels as documents, so that these verses could be searched for the level of relevance or similarity to the entered keywords. Furthermore, to determine the level of relevance between documents and keywords is calculated using the Vector Space Model. Based on the application that has been successfully built, the application can be show 10 documents related to the keywords that are searched and sorted from the most relevant, with the highest similarity value, namely 78.65%.

Keywords - Information Retrieval, Vector Space Model, Bible.

1. INTRODUCTION

Nowadays, it is undeniable that information has become something very important. Having a large number of documents is sometimes very annoying and troublesome, especially when we want to find the required documents with precise and fast information as necessary. Information can include documents, news, letters, stories, research reports, financial data, and others. The Documents consist of manual and digital documents, manual documents are documents that are in physical form and require a large storage area such as a room, while digital documents are electronic documents that are stored in electronic storage media. Currently, almost all of the physical document is converted into a digital document that can be stored in electronic storage media, making it easily accessible anywhere and anytime.

The Bible is the holy book for Christians, which consists of 66 books, namely 39 books of the Old Testament (OT) and 27 books of the New Testament (NT). In the NT consists of four main parts, that is the Gospel, the Book of History, the Apostolic Letters and the Book of Revelation. In the NT consists of four main parts, namely the Bible, the Book of History, the Apostolic Letters and Revelation. Gospel consists of four the Book is, Matthew, Mark, Luke, John Gospel contains the life, teachings, death, and resurrection of Jesus [1]. In Christian belief, the Gospel is called the Good News, because Christians believe that the narrative of the Gospels centered redemptive work of God to sinful mankind [2]. Many Christians want to study the gospel individually or collectively. However, they often experience problems in finding verses and topics in the Bible that are related to the needs the user wants to find, so they have to search individually each verse in the four Gospels to find the topic or verse they want to search for.

An information retrieval system is a system that has a function in finding the relevant information according to user requirements [3]. Information Retrieval is related to the general language text unstructured data or semi-structured. This causes additional challenges to the information retrieval system, which is complex and incomplete text structures, unclear and non-standard meanings, and the different languages as well as the inaccurate translations. Documents are unstructured information, the content of a document is highly dependent on the author of the document. Whereas currently many applications provide the search features using the data retrieval model. Data retrieval is related to data whose semantic structure is defined or clear when the user wants to find some data from the database by using a query that has been defined by each database, with the result that the output is a set of data that exactly matching with the query.

One of the mathematical models used in information retrieval systems is the concept of vector space model (VSM), where the user input and document translated into vectors. Then the vectors multiply operation performed and the results are used as a reference point to determine the relevance of the document you want to search based on the topics that the user entered [4]. In this study, we proposed to build an application of retrieval of information in the Bible, namely the Gospels (namely the Books of Matthew, Mark, Luke and John) using the vector space model method. Therefore, this application will be able to search the contents of the document in the form of verses in the Bible accurately and quickly based on the sentence entered. So that, can help user to discover the contents of the relevant documents in the Bible accordance with the necessary requirements.

In the rest of this paper, Section II discusses about the related works, and Section III presents our research method. Section IV reports our performance results and finally Section VI concludes the paper.

2. RELATED WORKS

Text mining on unstructured data supports the knowledge discovery process in large document collections. Text mining attempts to provide the solutions to the problems of information overload, processing, organization, or grouping and analyzing by using some techniques. In this study [7] the author using text mining to analyze the sentiments of people toward the candidates of a presidential election using the Naive Bayes algorithm. In this study, the opinion of people shared on social media as the documents, they used tweets are crawling from twitter.

Documents as data objects in the Information Retrieval System are sources of information. Documents are usually expressed in the form of an index or a keyword can be extracted directly from the text document or specifically defined in the indexing process that basically consists of a process of analysis and representation of the document. Information retrieval is a study in helping users to find information that suits their information needs [8]. Information retrieval relates to information retrieval whose content does not have a structure. Likewise, an expression of user requirements, called a query, also has no structure. This is what makes the difference information retrieval from database systems, documents are examples of unstructured information, where the contents of a document are highly dependent on the author of the document. The purpose of information retrieval is to meet the information needs of users to rediscover all the relevant documents, and at the same time rediscovering bit irrelevant documents [9].

As a system, the information retrieval system has several parts that make up the whole system. The process of information retrieval begins with building a database, namely document collection. Afterward, the database is built document searches can be performed, by entering a search query, then the system will perform text operations from the queries and document collections that exist in the database by selecting words or by removing the unnecessary parts in queries or document collections [10]. Furthermore, the system will perform queries and document formulations to calculate the level of similarity between the queries entered, with the available document collections, this stage uses certain algorithms or methods to determine similarity [12].

Vector Space Model (VSM) is a method to look at the degree of proximity or similarity term by the term weighting method. The document is seen as a vector that has distances and directions. In the vector space model, a term is represented by a dimension of vector space. The relevance of a document to a query based on the similarity between the query vector and document vectors [9]. Apart from that in VSM, the database of all the documents represented by the document term matrix or matrices term frequency, wherein each cell in the matrix corresponds to a given weight of a specified term in a document [11].

In this study, they build a full-text search engine to assist students in finding a thesis, a method of vector space model. The design of a full-text search engine using PHP and displayed in the form of web pages. Testing the relevance of the search engine is done by entering a few queries and the search engine will display the relevant documents [13]. Whereas research [14] aims to build a system that can help students find the linkages between theses in the library at the Universitas Ahmad Dahlan. The library at this university has no system that can determine the linkages or relationships between one thesis and another and there is also no system that can compare the conclusions and suggestions of the thesis. Therefore, in the Universitas Ahmad Dahlan library, there must be a system that can determine the linkages between the theses, making it easier to find linkages between these theses. In this study, the resulting system is made including input and editing of thesis data and determining the relationship between theses, and can display thesis data, namely displaying data suggestions, conclusions, contents, and thesis abstracts [14]. However, the document storage is not stored in the MySql database due to the indexing process and calculating the weight of the process takes a long time, this makes the limitation on the volume of the inputted document.

Furthermore, in [15] used the case study is to source digital library puppet. The purpose of this study is to improve the accuracy of search engines in searching for documents so that the search results for documents are closer to the needs of users. Search by inputting keywords using the vector space model and semantic search using domain ontology models. This search is by inputting keywords using a document index based on term weighting, while semantically searching uses a document index based on ontology domain metadata. The ontology domain is the extraction of knowledge from documents. In the semantic search, a choice of keyword combinations will appear on different domain concepts. This keyword combination is an extension of the keywords entered by the user. However, the user can only enter a two-word search query and the ontology domain that is created is still done manually.

Based on the shortcomings of several related studies, the author will build an information retrieval application in the Bible by providing a search feature to find relevant verses in the Bible. Users can enter more than five words in one sentence in the search query and the ranking is done automatically by the application which amounts to ten relevant documents. In addition, it will provide a feature to show the entire document of the relevant paragraph.

3. RESEARCH METHOD

In this section, the author discusses the research methodology to be used, namely Scrum. A scrum is a framework in the development of software, which is included in the well-known Agile methods. A Scrum is a framework that is adaptive, repeatable, fast, flexible, and effective which is designed to provide significant value quickly on a project. Scrum methodology is to apply the concept of a combination of iterative and incremental approach. A sprint is an iterative and as the sprint increases, the more features are implemented (incremental). The following are the five main phases of the scrum method [5, 6]:

Phase 1: Initiate, the initial phases of Scrum, which is to create the descriptions of the project to be created, which aims to form a team, compile a project vision as well as determine a product backlog.

Phase 2: Plan and Estimate, at this phase the aim is to plan in preparation, sprint execution, writing of the user stories, elaborating tasks for each user story, estimating the value of each user story and task.

Phase 3: Implement, the aim of this stage is the execution of each task or phase that has been analyzed and designed, as well as undertake activities to form the product. This stage also includes a daily standup meeting, which is the stage for evaluating the products that have been built.

Phase 4: Review and Retrospect, in conducting a review of the work that has been done in the form of testing the project, so that it can be assessed whether the project is in conformity with the planned and whether the project has eligible for the requirements, this is done in this phase.

Phase 5: Release, this phase aims to conclude the results that have been achieved, i.e, the application made is in accordance with the previous stages, then the application that has met the criteria and is ready for use by the user.

4. PERFORMANCE RESULT

4.1. Implementation

In this section, we will discuss the application that was built, namely the application of retrieval of information in the Bible using the vector space model method. This application is expected to assist Christians in finding the contents of documents in the Bible in accordance with the necessary requirements. This application uses the search feature, where users can enter a keyword currently in the form of words or sentences, then the application will perform the search process by performing calculations using the VSM method. Resulting in, the application will be shown the verses in the Bible that contain words or sentences that have been entered by the user. The paragraphs shown are based on rank, in the form of ten documents that have been sorted from the most relevant to the least relevant. Figure 1 shows an overview of the content outline of the information retrieval application in the Bible.

The process begins with building a database that is document collection, in this application the verses in the Bible are document collections. After the database is built, the document search process can be run. When the user enters a search query in the form of a sentence or topic that he wants to search for the paragraph, then the system will perform text operations from the entered query and document collections in the database. Furthermore, the system will perform queries and document formulations to calculate the degree of similarity between the queries captured and the available document collections. At this stage, we use the VSM method to determine the level of similarity. Hereinafter, calculating the degrees of similarity between the query and document, then the system will give the weight to the query and document, with the result that it can be seen which document is the most suitable for the query. Thereupon, the document where the highest weighting the ratings given so the system can show the which documents are most relevant to the search query.

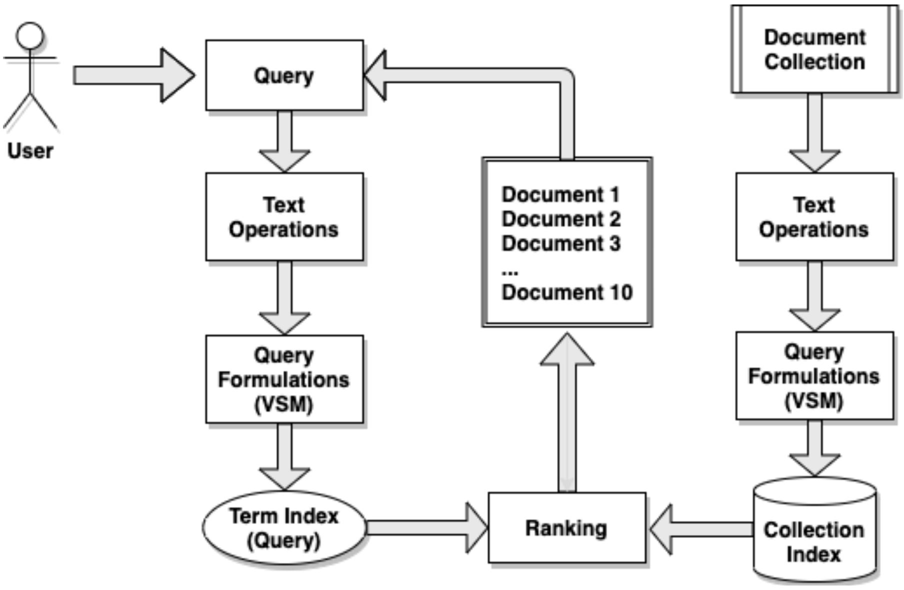


Figure 1 Content Outline of the Application

The following is the process to obtain the results of a query is entered on the application of information retrieval in the Bible by performing the calculations using the VSM methods. Firstly, we discard the prepositions and conjunctions. For example, there are 3 documents *{D=D$_1$, D$_2$, D$_3$}* and a query (*Q*), as follows:

D$_1$: *"Ketika Yesus mengangkat muka-Nya, Ia melihat orang-orang kaya memasukkan persembahan mereka ke dalam peti persembahan. Ia melihat juga seorang janda miskin memasukkan dua peser ke dalam peti itu."*
D$_2$: *"Lalu Ia berkata: Aku berkata kepadamu sesungguhnya, janda ini memberi lebih banyak daripada semua orang itu."*
D$_3$: *"Sebab mereka semua memberi persembahannya dari kelimpahannya, tetapi janda ini memberi dari kekurangannya, bahkan ia memberi seluruh nafkahnya."*
Q : *"Pemberian Janda Miskin."*

Afterwards, using stemmer in document collection and queries, which is an application that is used to remove the prepositions and conjunctions, and remove affixes (prefix and suffix). Therefore, the 3 documents and a query become:

D$_1$: *Yesus angkat muka, Ia lihat orang-orang kaya masuk sembah mereka dalam peti sembah. Ia lihat orang janda miskin masuk dua peser dalam peti.*
D$_2$: *Ia kata Aku kata pada sungguh, janda beri lebih banyak pada semua orang.*
D$_3$: *Mereka semua beri sembah limpah, janda beri kurang, ia beri seluruh nafkah.*
Q : *Beri Janda Miskin.*

Furthermore, define the minterm to determine the possible word frequency patterns. The minterm length is based on the number of words entered in the query. Hereinafter, it is converted into an orthogonal vector based on the emerging minterm pattern. Then, the index term is calculated by using Equation 1. Meanwhile, the correlation factor can be calculated using Equation 2. Where, $\overrightarrow{k_i}$ is an index term to-*i,* the orthogonal vector according to the minterm pattern used is denoted by $\overrightarrow{m_r}$, while $C_{i,r}$ for the correlation factor between index term *i* with minterm *r*. The index term *i* weight in document j is denoted by $W_{i,j}$ and $g_i(m_r)$ is the weight index term *k$_i$*

in minterm m_r, whilst $\overrightarrow{d_j}$ is the j^{th} document vector and $\vec{q}$ is the vector query, as well as the weight index term in query i is q_i.

$$\overrightarrow{k_\iota} = \frac{\sum_{\forall r, g_i(m_r)=1}^{n} c_{i,r} * \overrightarrow{m_r}}{\sqrt{\sum_{\forall r, g_i(m_r)=1}^{n} {c_{i,r}}^2}} \tag{1}$$

$$C_{i,r} = \sum_{d_j | g_i(\overrightarrow{d_j}) = g_i(m_r)}^{n} W_{i,j} \tag{2}$$

$$\overrightarrow{d_J} = \sum_{i=1}^{n} W_{i,j} * \overrightarrow{k_\iota} \qquad \vec{q} = \sum_{i=1}^{n} q_i * \overrightarrow{k_\iota} \tag{3}$$

Moreover, converted documents and queries into vectors using Equation 3, then sort the documents based on similarity, by calculating the multiplication of vectors using the following equation:

$$sim(\vec{q}) = \frac{\overrightarrow{d_J} * \vec{q}}{|\overrightarrow{d_J}|\,|\vec{q}|} \tag{4}$$

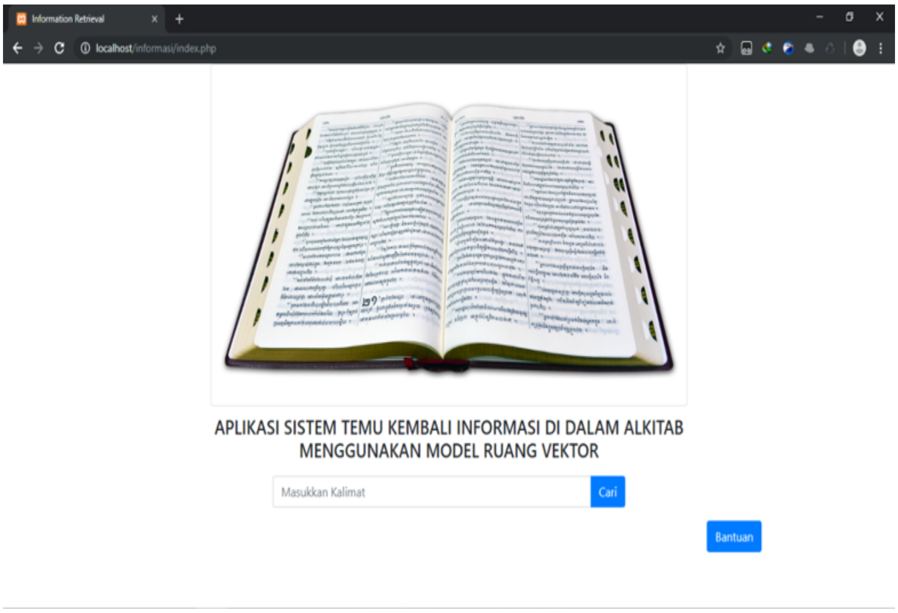
In this case, based on the query the minterm used is m_1, m_2, m_3, as shown in Table 1. Thereafter, calculate the frequency of words in the collection of documents that match the query and determine the orthogonal vector according to the minterm used. The frequency of the words in the document collection, that the correlation of each term is obtained, that is $C_{1,1} = 0$, $C_{1,2} = 1$, $C_{1,3} = 3$, $C_{2,1} = 1$, $C_{2,2} = 1$, $C_{2,3} = 1$, $C_{3,1} = 1$, $C_{3,2} = 0$, *and* $C_{3,3} = 0$.

Tabel 1 Word Frequency in Document Collection

	Beri	*Janda*	*Miskin*	**Vektor orthogonal**
$\boldsymbol{D_1}$	0	1	1	$\overrightarrow{m_1}$
$\boldsymbol{D_2}$	1	1	0	$\overrightarrow{m_2}$
$\boldsymbol{D_3}$	3	1	0	$\overrightarrow{m_3}$
$\boldsymbol{Q}$	1	1	1	-

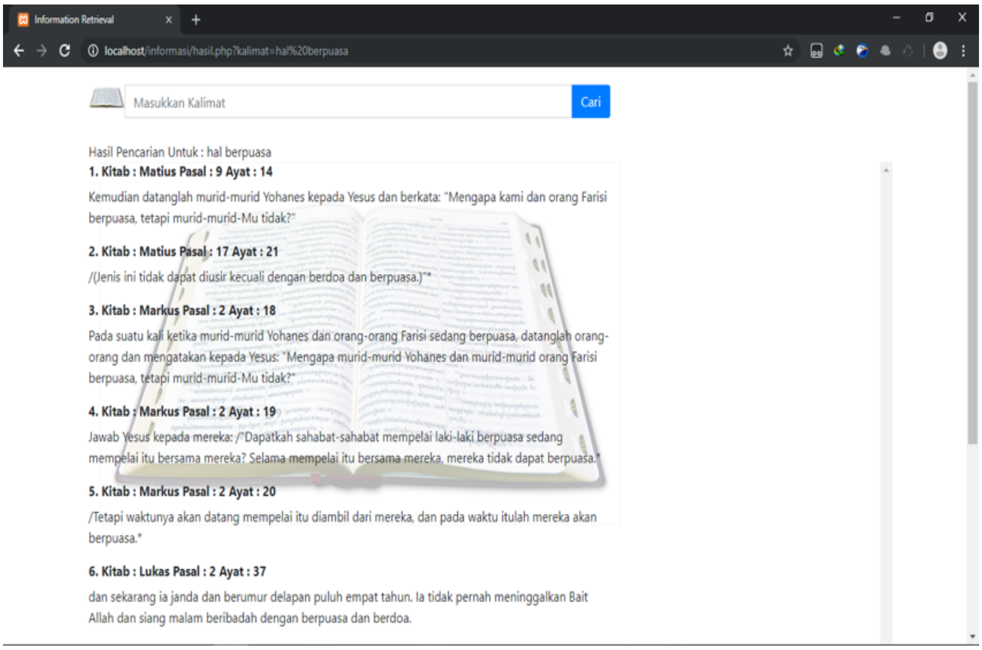
Then the index-term are obtained as follows: $\overrightarrow{k_1} = \frac{\overrightarrow{m_2} + 3\overrightarrow{m_3}}{\sqrt{10}}$, $\overrightarrow{k_2} = \frac{\overrightarrow{m_1} + \overrightarrow{m_2} + \overrightarrow{m_3}}{\sqrt{3}}$, $\overrightarrow{k_3} = \overrightarrow{m_1}/\sqrt{1}$. Further, converted the query document into vector form $\overrightarrow{d_1} = \overrightarrow{k_2} + \overrightarrow{k_3} = 1{,}5773\,\overrightarrow{m_1} + 0{,}5773\,\overrightarrow{m_2} + 0{,}5773\,\overrightarrow{m_3}$, $\overrightarrow{d_2} = \overrightarrow{k_1} + \overrightarrow{k_2} = 0{,}5773\,\overrightarrow{m_1} + 0{,}8935\,\overrightarrow{m_2} + 1{,}5259\,\overrightarrow{m_3}$, $\overrightarrow{d_3} = 3\overrightarrow{k_1} + \overrightarrow{k_2} = 0{,}5773\,\overrightarrow{m_1} + 1{,}5259\,\overrightarrow{m_2} + 3{,}4233\,\overrightarrow{m_3}$ *and* $\overrightarrow{q_1} = \overrightarrow{k_1} + \overrightarrow{k_2} + \overrightarrow{k_3} = 1{,}5773\,\overrightarrow{m_1} + 0{,}8935\,\overrightarrow{m_2} + 1{,}5259\,\overrightarrow{m_3}$. Hereafter, the similarity of documents and the query $Sim\,(\overrightarrow{d_1}, \vec{q}) = 0.9230$, $Sim\,(\overrightarrow{d_2}, \vec{q}) = 0.9177$, and $Sim(\overrightarrow{d_3}, \vec{q}) = 0.8352$. Based on the results of the similarity calculation, it can be seen that for the ranking order, documents that are closed to 1 are the documents that are ranked first. Therefore, in this case the first document is the most relevant document to the query with a value of 0.9230, followed by the second document and the third document.

In this application, we use the Bible in Bahasa Indonesia, with the main menu of applications that have been built are shown in Figure 2, where there is a search field that can be filled by the user in the form of keywords that will be searched. Furthermore, Figure 3 shows the display of search results in applications where keywords entered by the user, the related document will be displayed on this page.



APLIKASI SISTEM TEMU KEMBALI INFORMASI DI DALAM ALKITAB MENGGUNAKAN MODEL RUANG VEKTOR

Figure 2 The Interface of Main Menu



Hasil Pencarian Untuk : hal berpuasa

1. Kitab : Matius Pasal : 9 Ayat : 14

Kemudian datanglah murid-murid Yohanes kepada Yesus dan berkata: "Mengapa kami dan orang Farisi berpuasa, tetapi murid-murid-Mu tidak?"

2. Kitab : Matius Pasal : 17 Ayat : 21

/(Jenis ini tidak dapat diusir kecuali dengan berdoa dan berpuasa.)"*

3. Kitab : Markus Pasal : 2 Ayat : 18

Pada suatu kali ketika murid-murid Yohanes dan orang-orang Farisi sedang berpuasa, datanglah orang-orang dan mengatakan kepada Yesus: "Mengapa murid-murid Yohanes dan murid-murid orang Farisi berpuasa, tetapi murid-murid-Mu tidak?"

4. Kitab : Markus Pasal : 2 Ayat : 19

Jawab Yesus kepada mereka: /"Dapatkah sahabat-sahabat mempelai laki-laki berpuasa sedang mempelai itu bersama mereka? Selama mempelai itu bersama mereka, mereka tidak dapat berpuasa.*

5. Kitab : Markus Pasal : 2 Ayat : 20

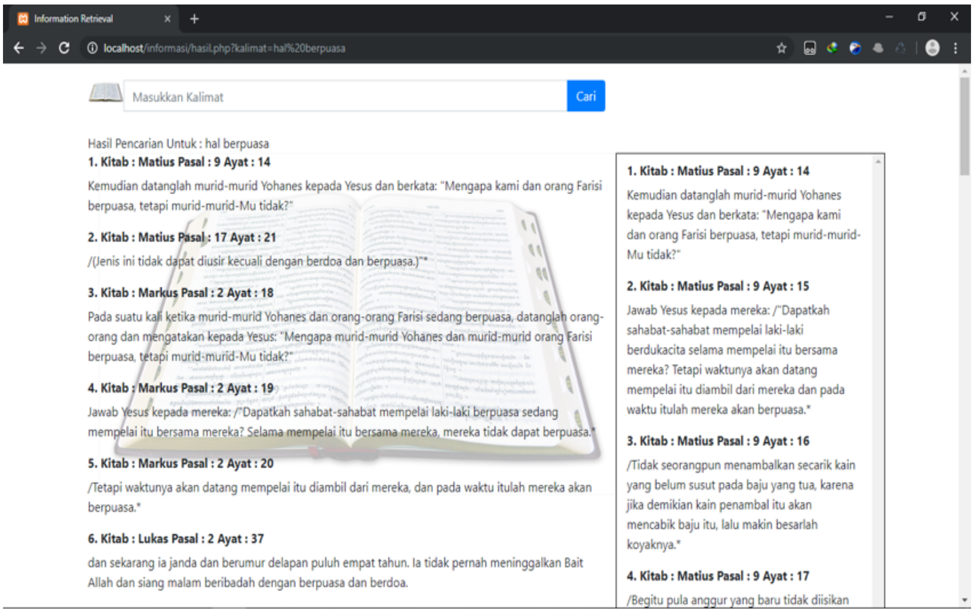
/Tetapi waktunya akan datang mempelai itu diambil dari mereka, dan pada waktu itulah mereka akan berpuasa.*

6. Kitab : Lukas Pasal : 2 Ayat : 37

dan sekarang ia janda dan berumur delapan puluh empat tahun. Ia tidak pernah meninggalkan Bait Allah dan siang malam beribadah dengan berpuasa dan berdoa.

Figure 3 The Application Interface of Search Results

Whereas Figure 4 shows the search results from the main menu display in the application, and on the right side there are all the verses related to the search results starting from the verse obtained by the search results. In addition, the search results if the relevant verse is less than ten verses are shown in Figure 5.



Hasil Pencarian Untuk : hal berpuasa

1. Kitab : Matius Pasal : 9 Ayat : 14

Kemudian datanglah murid-murid Yohanes kepada Yesus dan berkata: "Mengapa kami dan orang Farisi berpuasa, tetapi murid-murid-Mu tidak?"

2. Kitab : Matius Pasal : 17 Ayat : 21

/(Jenis ini tidak dapat diusir kecuali dengan berdoa dan berpuasa.)"*

3. Kitab : Markus Pasal : 2 Ayat : 18

Pada suatu kali ketika murid-murid Yohanes dan orang-orang Farisi sedang berpuasa, datanglah orang-orang dan mengatakan kepada Yesus: "Mengapa murid-murid Yohanes dan murid-murid orang Farisi berpuasa, tetapi murid-murid-Mu tidak?"

4. Kitab : Markus Pasal : 2 Ayat : 19

Jawab Yesus kepada mereka: /"Dapatkah sahabat-sahabat mempelai laki-laki berpuasa sedang mempelai itu bersama mereka? Selama mempelai itu bersama mereka, mereka tidak dapat berpuasa.*

5. Kitab : Markus Pasal : 2 Ayat : 20

/Tetapi waktunya akan datang mempelai itu diambil dari mereka, dan pada waktu itulah mereka akan berpuasa.*

6. Kitab : Lukas Pasal : 2 Ayat : 37

dan sekarang ia janda dan berumur delapan puluh empat tahun. Ia tidak pernah meninggalkan Bait Allah dan siang malam beribadah dengan berpuasa dan berdoa.

1. Kitab : Matius Pasal : 9 Ayat : 14

Kemudian datanglah murid-murid Yohanes kepada Yesus dan berkata: "Mengapa kami dan orang Farisi berpuasa, tetapi murid-murid-Mu tidak?"

2. Kitab : Matius Pasal : 9 Ayat : 15

Jawab Yesus kepada mereka: /"Dapatkah sahabat-sahabat mempelai laki-laki berdukacita selama mempelai itu bersama mereka? Tetapi waktunya akan datang mempelai itu diambil dari mereka dan pada waktu itulah mereka akan berpuasa.*

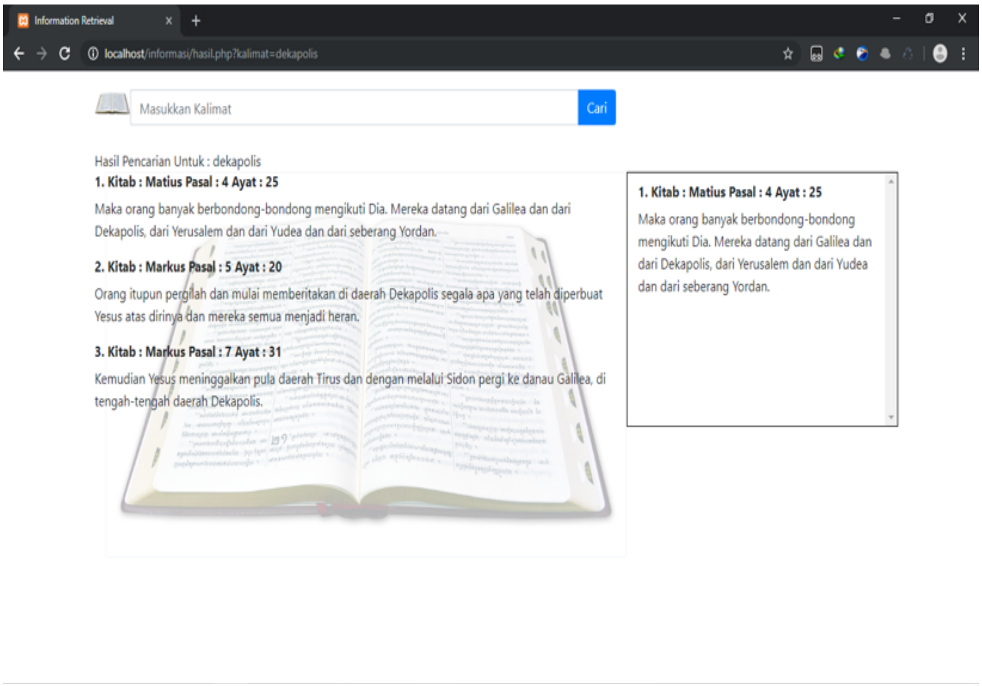
3. Kitab : Matius Pasal : 9 Ayat : 16

/Tidak seorangpun menambalkan secarik kain yang belum susut pada baju yang tua, karena jika demikian kain penambal itu akan mencabik baju itu, lalu makin besarlah koyaknya.*

4. Kitab : Matius Pasal : 9 Ayat : 17

/Begitu pula anggur yang baru tidak diisikan

Figure 4 The Interface of Application Results with the Relevant Documents.



Hasil Pencarian Untuk : dekapolis

1. Kitab : Matius Pasal : 4 Ayat : 25

Maka orang banyak berbondong-bondong mengikuti Dia. Mereka datang dari Galilea dan dari Dekapolis, dari Yerusalem dan dari Yudea dan dari seberang Yordan.

2. Kitab : Markus Pasal : 5 Ayat : 20

Orang itupun pergilah dan mulai memberitakan di daerah Dekapolis segala apa yang telah diperbuat Yesus atas dirinya dan mereka semua menjadi heran.

3. Kitab : Markus Pasal : 7 Ayat : 31

Kemudian Yesus meninggalkan pula daerah Tirus dan dengan melalui Sidon pergi ke danau Galilea, di tengah-tengah daerah Dekapolis.

1. Kitab : Matius Pasal : 4 Ayat : 25

Maka orang banyak berbondong-bondong mengikuti Dia. Mereka datang dari Galilea dan dari Dekapolis, dari Yerusalem dan dari Yudea dan dari seberang Yordan.

Figure 5 The Interface of Application Results Less Than Ten Relevant Documents

4.2. Testing

In this section, we will discuss the test objectives of applications that have been built, that is application testing done to find out the conditions or requirements of the user and the system running features and functions, as well as to know the reaction of the application with input provided by the user. The following are test cases in the form of checkpoints from the information retrieval application in the Bible using a vector space model that has been built, among others: the application can rank documents based on the keyword searched and displays relevant verses or documents with a maximum of ten documents. In addition, the application can run on various commonly used web browsers with the appearance of an attractive application.

Based on the results of the tests that have been done, the application can be shown the verses are related to the search keywords entered. Applications can also sort from most relevant to irrelevant. In addition, it can also display the complete paragraph or the next paragraph of the relevant document from the keyword. The application interface is also quite interesting for users that can run on several web browsers, such as Google Chrome Version 75.0.3770.142 (64 bit), Mozilla Firefox Version 67.0.4 (64-bit), and Microsoft Edge Version 42.17134.1.0 (64-bit). All the features of this application can also be run properly.

5. CONCLUSION

Based on the research that has been done, it can be concluded that the application of the information retrieval system in the Bible using the vector space model has been successfully implemented in the process of searching for verses in the Bible (Gospel). The application built can provide information and convenience in searching for documents or verses in the Bible, namely the Bible, based on keywords entered in the form of sentences or topics. Furthermore, the search results will be sorted based on the ranking of the document weight according to the user input keyword with a similarity value of 78.65%.

For future work, expected applications can be developed by combining the optimization algorithm, so that results can be obtained more quickly. In addition, developed by adding the Book of Acts, Letters, and Revelation Apostolic.

REFERENCES

[1] Lori., 2019, Orang Kristen Wajib Tahu Pembagian Kitab Alkitab ini, https://www.jawaban.com/read/article/id/2016/10/28%2008:00:00/58/161027163328/orang_kristen_wajib%20_tahu_pembagian_kitab_alkitab_ini. diakses 28 October 2019.

[2] Suprandono, Y.R., Keyakinan Iman Kita: Alkitab adalah Firman Allah, *Artikel Teologi,* https://sttkharisma.ac.id/keyakinan-iman-kita-alkitab-adalah-firman-allah.html

[3] Rahmah, E., Akses dan Layanan Perpustakaan Teori dan Aplikasi, Jakarta: Prenadamedia Group, 2018.

[4] Bunyamin, H., Negara, C,P., "Aplikasi Information Retrieval (IR) CATA Dengan Metode Generalized Vector Space Model," Jurusan Teknik Informatika Fakultas Teknologi Informasi, Universitas Kristen Maranatha, vol. IV, no. 1, p. 29- 38, 2018.

[5] Schwaber.K., Sutherland.J., “Panduan Scrum” Scrum.org and ScrumInc, 2014.

[6] Satpathy, T., A Guide to the Scrum Body Of Knowledge (SBOKTM Guide), Arizona: SCRUMstudyTM, a brand of VMEdu, Inc., 2016.

[7] Wongkar.M., Angdresey.A., “Sentiment Analysis Using Naïve Bayes Algorithm of The Data Crawler: Twitter”. *In 2019 Fourth International Conference on Informatics and Computing (ICIC),* pp. 1-5, 2019.

[8] Chu, H., Information and Representaion and Retrieval in the Digital Age, New Jersey: Information Today, 2009.

[9] Abdillah, A,A., Muktyas, I,B., “Implementasi Vector Space Model Untuk Pencarian Dokumen,” Seminar Nasional Matematika dan Pendidikan Matematika, Malang, 2013.

[10] Amin, F., “Implementasi Search Engine (Mesin Pencari) Menggunakan Metode Vector Space Model,” 1127-1627-1-PB, P. 14, 2011.

[11] Waranu, I,P,A., “Implementasi Metode Generalized Vector Space Model Pada Information Retrieval System,” Makalah-Stki1204505042, P. 21, 2015.

[12] Amin, F., “Sistem Temu Kembali Informasi dengan Pemeringkatan Metode Vector Space Model”, Jurnal Teknologi Informasi DINAMIK Vol.18, No.2, Juli 2013.

[13] Kusuma, M,A., Kamayani, M., Avorizano, A., "Pencarian Full Text Pada Koleksi Skripsi Fakultas Teknik Uhamka Menggunakan Metode Vector Space Model," Seminar Nasional Teknoka, vol. II, no. 2, pp. 96-102, 2017.

[14] Mulyadin, Aribowo, E., "Sistem Penentuan Keterkaitan Antar Skripsi Bersadarkan Keyword Seeking," Jurnal Sarjana Teknik Informatika, vol. II, no. 1, pp. 856-865, 2014.

[15] Hadhiatma, A., "Pencarian Dokumen Berdasarkan Kombinasi Antara Model Ruang Vektor Dan Model Domain Ontologi," Seminar Nasional Informatika, pp. 111-117, 2010.