# Automatic Text Summarization using Maximum Marginal Relevance for Health Ethics Protocol Document in Bahasa

**Doni Putra Purbawa**
*Informatics Department, Faculty of Intelligent Electrical and Informatics Technology*
*Institut Teknologi Sepuluh Nopember (ITS)*
Surabaya, Indonesia
doniputrapurbawa@gmail.com

**Malikhah**
*Informatics Department, Faculty of Intelligent Electrical and Informatics Technology*
*Institut Teknologi Sepuluh Nopember (ITS)*
Surabaya, Indonesia
meli2511@gmail.com

**Ratih Nur Esti Anggraini**
*Informatics Department, Faculty of Intelligent Electrical and Informatics Technology*
*Institut Teknologi Sepuluh Nopember (ITS)*
Surabaya, Indonesia
ratih_nea@if.its.ac.id

**Riyanarto Sarno**
*Informatics Department, Faculty of Intelligent Electrical and Informatics Technology*
*Institut Teknologi Sepuluh Nopember (ITS)*
Surabaya, Indonesia
riyanarto@gmail.com

***Abstract*— The rapid development of science and technology in the health sector cannot be separated from the support of health research results. Before the research with human as a subject is conducted, every health research in Indonesia is required to make a health ethics protocol document in Bahasa which must comply with basic ethical principles. To determine whether the health research ethics protocol document has met the ethical principles, the health research ethics protocol document will be reviewed by a competent reviewer. The health research ethics protocol document consists of several parts and has a large number of pages, so to conduct a review, reviewers need a long time to understand and analyze the health research ethics protocol document. To reduce the review time, an automatic text summarization (ATS) is needed. ATS extracts important information in health research ethics protocol documents and presents it to reviewers. This research uses cosine similarity and Maximum Marginal Relevance (MMR) and TextRank to summarize the document. The MMR method is considered to have more stable results than TextRank based on the ROUGE evaluation results. The evaluation of result with the ROUGE Toolkit showed F-score value of 19.92% for document 1 and 10.98% for document 2 using MMR.**

***Keywords—automatic text summarization, bahasa, ethics protocol document, MMR, health***

## I. INTRODUCTION

The rapid development of science and technology in the health sector cannot be separated from the support of health research results [1]. Before the research results can be used safely and effectively, the research must include humans and animals as research subjects. The ethical justification for conducting health research involving humans and animals is from social and scientific value, because the purpose of this health research is to generate the knowledge and tools necessary to protect and improve public health. Health professionals, patients, researchers, the makers of policy, public health officials, pharmaceutical firm and others rely on the results of health research for activities and decisions that impact individual and community health, well-being, and use of limited resources. Consequently, the proposed research must be scientific after careful inspection by researchers, sponsors, research ethics committees, and health authorities. The proposed research must also build on an adequate prior knowledge base, and are tend to produce valuable and important information [2].

The implementation of research in the health sector itself must be carried out in accordance with ethical standards for health research [3]. In Indonesia itself, before conducting health research, researchers are required to make a health ethics protocol document written in Bahasa which is prepared based on a predetermined format. The ethical protocol document will then be reviewed by a competent person in the health sector to ensure that the ethical protocol document has complied with the rules and does not violate the provisions that have been made previously. Based on the guidelines and ethical standards of national health research and development from the Indonesian Ministry of Health [4], reviewers of health ethics protocols must also ensure that the research has complied with three basic ethical principles, (1) the principle of respecting human dignity as individuals who have freedom of will or choice and are at the same time personally responsible for their own decisions; (2) the principle of doing good and not harming. The ethical principle of doing good concerning the obligation to help others is carried out by seeking maximum benefits and minimizing losses; (3) the principle of justice, refers to the ethical obligation to treat everyone equally and not to discriminate.

The health research ethics protocol document itself consists of several parts and has many pages. It is necessary to summarize the health research ethics protocol document so that the reviewers does not need to read and analyze the whole document which take sufficient time but is still able to take the essence of the health research ethics protocol documents. The health research ethics protocol documents which summarized with an automatic text summarization will reduce the time needed in the process of reviewing the document. This summary of the health research ethics protocol document will take the important parts of the document. Automatic text summarization can help extract important information in



978-1-6654-4059-2/21/$31.00 ©2021 IEEE

health research ethics protocol documents and present it to reviewers. With the summary, the time needed to review health research ethics protocol documents will be reduced.

Research to summarize Indonesian language text documents has been carried out before, but research that focuses on summarizing health research ethical protocol documents has never been done before. Therefore, the aim of this research is generating automatic text summarization of health research ethics protocol document to help reviewers to understand the contents of the document written in Bahasa and reduce the time needed to conduct a review. This research proposed cosine similarity with MMR to extract summarization from health research ethical protocol documents.

## II. Related Work

Nowadays, with the increasing number of textual content that grows exponentially on the internet and various articles, scientific papers, legal documents, etc., Automatic Text Summarization (ATS) is becoming increasingly important [5]. Summarizing text manually takes a lot of time, effort, cost, and even becomes impractical with a lot of textual content. ATS itself can be performed to summarize both single and multi-document [6]. Research related to ATS in Bahasa has also been done before.

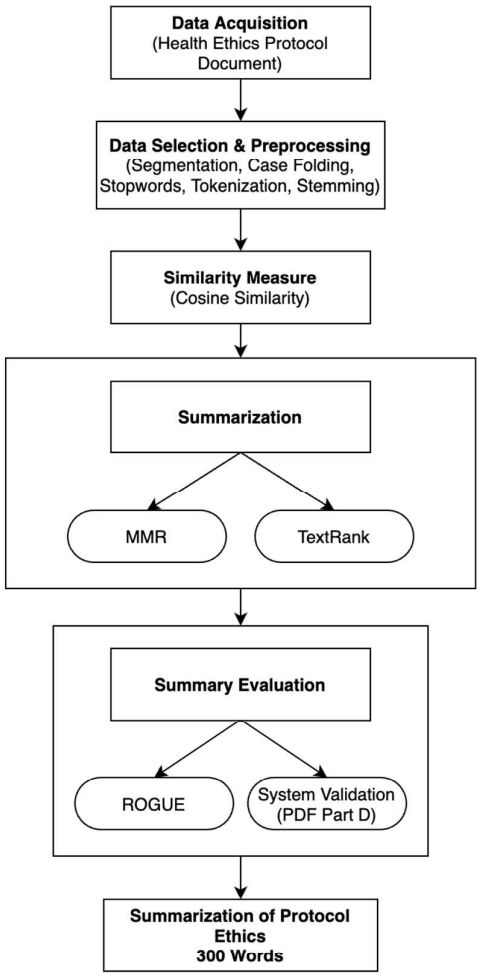


Fig. 1. Research Methodology

There are several recent studies related to Indonesian language ATS, among others, in 2017, Sabuna et al [7] who summarized 50 news text as training data. This study conducted in four steps. The first step was collecting document that will be used as data training and data testing. The second step was producing a model using decision tree method. In the training step, this study used pre-processing like tokenization, stemming, and stopword. TF-IDF is used in this study to score the sentence. Each sentence weight will be used as the training data of decision tree algorithm to produce decision model. After producing a decision tree model, the next step was testing the model to produce summarization system. The last step of this study was evaluation, to compare the summarization produced by the model and the manual summary by users. By using sentence scoring method and decision tree on Indonesian text summarization, this study gets the f-score of 0,58, precision average of 0,54, and recall of 0,66.

Gunawan et al [8] researched to summarize thirty online news articles using TextRank and Maximum marginal relevance (MMR) methods. The news articles are pre-processed to produce clean text. After that, this research used the TextRank method to extract the essential sentences using the similarity measurement. This process produced a text that has been summarized. However, the digested text is still containing similar sentences. Therefore, this research carries out the following procedure: calculating Maximal Marginal Relevance (MMR) to reduce similar sentences. The result of this process is the final text summary. The evaluation used ROUGE-1 and ROUGE-2 with the average F-score is 0.5103 and 0.4257, respectively.

The latest research in 2021 by Putra et al conducted a study to make a summary of student essay assigment [9]. This study used sentence weighting feature to select the sentence with the highest weight in one cluster. This method helped the system quickly get the important information from the document. The results of this study indicate that the system succeeds in summarizing text with an average evaluation of the values of precision, recall, accuracy, and F-score of 0.52, 0.54, 0.70, and 0.52, respectively.

## III. Proposed Method

This research produced a text summary containing a maximum of 300 words from two health research ethics protocol documents. The type of document that will be summarized is a single document. The process of making a text summary is divided into 5 steps, namely data acquisition, data selection and pre-processing, calculate similarity measure, summarization, and summary evaluation.

The methodology of this research as shown in Fig. 1 are explained as follows:

1. Data acquisition. In this step we got the docment to be summarized in the form of health research ethics protocol documents
2. Data selection and data pre-processing. There are 5 things that are done in the pre-processing stage, namely segmentation, case folding, stopwords, tokenization, and stemming.
    a. Content extraction. At this stage, the sentences contained in the ethical protocol document are extracted using Regular Expression (RegEx). The ethics protocol document has sections from A to Z.

Where A – C is the identity of the author or respondent, then D is a summary written by the respondent, F is the reference of ethical protocol document, and E, G-Z is the important points related to the Health Ethics Protocol that are interconnected.

b. Case folding. At this stage all sentences are converted to lowercase.

c. Stopwords removal. Because the document to be summarized is documents written in Bahasa, the stopword used is the Indonesian or Bahasa stopword. At the stopword stage, meaningless words such as connecting words will be removed.

d. Tokenization. At this stage each sentence is broken down into words or phrases that stand alone and are separated by spaces

e. Stemming. At the stemming stage, the 'sastrawi' library is used. the 'sastrawi' library is a python library which helps convert words into basic word forms based on Indonesian or Bahasa rules.

3. Calculation of similarity [10]. At this stage, the similarity between sentences with other sentences will be calculated using cosine similarity. The equation 1 is used to calculate the cosine similarity between 2 sentences.

$$Cosine\ Similarity = \frac{\sum_{n=1}^{j}(nA\ x\ nB)}{\sqrt{\sum_{n=1}^{j}(nA)^2}\sqrt{\sum_{n=1}^{j}(nB)^2}} \quad (1)$$

where, $nA$ is the number of occurrences of the nth index word from the word list in sentence $A$, while $nB$ is the number of occurrences of the nth index word from the word list in sentence $B$.

4. Summarization. this study uses two methods for summarization, namely MMR and TextRank. MMR [8] is a simple and effective extraction method to reduce duplication in text summarization results. The possibility of duplication or redundancy will be greater when a text summary has a high level of similarity between sentences. Therefore, methods to eliminate redundancy are indispensable. Equation 2 is used to calculate MMR so that redundancy can be reduced.

$$MMR\ (S_i) = \lambda \times Sim_1(S_i, D) - (1 - \lambda) \times Sim_2(S_i, Sum) \quad (2)$$

Where, $Sum$ contains sentences that have been extracted into summary, $S$ are sentences that have not been extracted. $Sim$ is the value of cosine similarity between two sentences. $\lambda$ aims to adjust the given value to emphasize relevance and avoid redundancy. The value of $\lambda$ ranges between 0 and 1. The higher the MMR value indicates that the sentence has low similarity. The advantage of this MMR method is that MMR has the ability to generate sentences with new information obtained from the above equation. TextRank is a graph-based ranking algorithm for text summarization [11]. TextRank algorithm performs sentence ranking based on content overlap similarity. Then the sentences that often appear in document will be included in the summary.

5. Summary Evaluation

This study used ROUGE (Recall Oriented Understudy for Gisting Evaluation) [12] to evaluate the summary generated by proposed method with the original summary. There are several metrics in ROUGE, namely ROUGE-1, ROUGE-2, ROUGE-3, and so on refers to the overlap of unigram (each word) between the system and reference summaries. This study mainly focuses on improving ROUGE-2, which refers to the overlap of bigrams between the system and reference summaries. ROUGE-L Longest Common Subsequence (LCS) based statistics. Longest common subsequence problem takes into account sentence level structure similarity naturally and identifies longest co-occurring in sequence n-grams automatically. ROUGE-W is a Weighted LCS, this method develops the ROUGE-L way by looking at the sentences in the candidate summary as a series of words that will later be matched in the references summary using dynamic programming [13].

## IV. Result and Discussion

This study used data on health research ethics protocols documents that have been filled out by two respondents. First protocol document entitled "Studi Kasus Analisis Faktor Kepatuhan Karyawan" and second protocol document entitled

TABLE I. ROUGE Score for Extracted Data

| Data | ROUGE Score | | | | |
|---|---|---|---|---|---|
| | 2 | 3 | 4 | L | W |
| Doc 1 | P: 27.27<br>R: 27.00<br>F: 27.14 | P: 25.68<br>R: 25.42<br>F: 25.55 | P: 24.75<br>R: 24.50<br>F: 24.62 | P: 30.54<br>R: 30.23<br>F: 30.38 | P: 30.54<br>R: 30.23<br>F: 30.38 |
| Doc 2 | P: 59.06<br>R: 58.86<br>F: 58.96 | P: 54.21<br>R: 54.03<br>F: 54.12 | P: 50.34<br>R: 50.17<br>F: 50.25 | P: 54.18<br>R: 54.00<br>F: 54.09 | P: 54.18<br>R: 54.00<br>F: 54.09 |

TABLE II. Pre-processing Step

| Process | Result |
|---|---|
| Raw Data | Kriteria exlusi adalah jajaran manajemen, administrasi perkantoran dan petugas sim RS penelitian terhadap responden dihentikan apabila resonden tidak setuju untuk mengisi Petugas rumah sakit yang terlibat sebagai calon subyek atau responden penelitian akan mendapatkan penjelasan tentang pelaksanaan penelitian. |
| Case Folding (Lower case) | kriteria exlusi adalah jajaran manajemen, administrasi perkantoran dan petugas sim rs penelitian terhadap responden dihentikan apabila resonden tidak setuju untuk mengisi petugas rumah sakit yang terlibat sebagai calon subyek atau responden penelitian akan mendapatkan penjelasan tentang pelaksanaan penelitian. |
| Tokenization | ['kriteria','exlusi','adalah', 'jajaran','manajemen', 'administrasi', 'perkantoran','dan', 'petugas','sim','rs','tidak', 'relevan','tidak','ada', 'intervensi','penelitian', 'terhadap','responden','dihentikan', 'apabila','resonden','tidak', 'setuju','untuk','mengisi','lembar', 'kuisioner'] |
| Stopwords Removal | ['kriteria','exlusi','jajaran', 'manajemen','administrasi', 'perkantoran','petugas','sim','rs', 'relevan','intervensi','penelitian', 'responden','dihentikan','resonden', 'setuju','mengisi','lembar', 'kuisioner] |
| Stemming | ['kriteria','exlusi','jajar', 'manajemen','administrasi','kantor', 'tugas','sim','rs','relevan', 'intervensi','teliti','responden', 'henti','resonden','tuju','isi', 'lembar','kuisioner'] |


Fig. 2. Word Cloud From Health Research Ethics Protocols Document 1

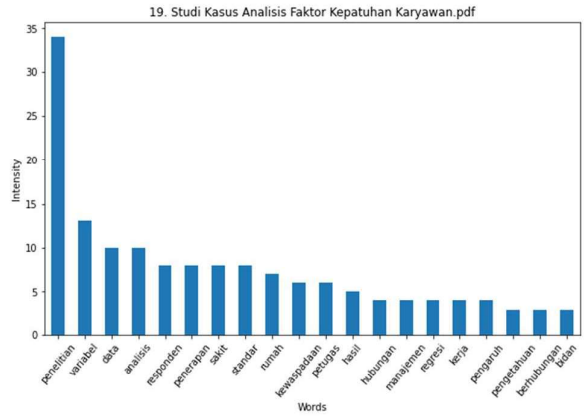


Fig. 3. Frequency of Each Word from Health Research Ethics Protocols Document 1

"Studi Kasus Asuhan Kefarmasian". Data from points D, G to Z were extracted using RegEx to retrieve only the important parts from the documents. In this process, this study used two RegEx commands "(\n{1,})\s*" to replace all characters "\n" with spaces " " and "\d{1,}\.[^*]*\ *" to omit the title or question for respondents (retrieving content only). Then each data extraction point is combined into one paragraph in the dataframe. Part D is saved as the summary of the document by the respondent. Part G until Z is combined into one. This part will be summarized using the proposed method.

Part G to part Z which has been extracted with RegEx then calculated the similarity with the original health ethics protocol document. The results show that the extraction from RegEx gives results that are almost similar to the original ethical protocol documents, where for the first ethical protocol document, the average similarity of RegEx extraction result with the original document is 86.51%, while for the second ethical protocol document, the average similarity of RegEx extraction result with the original document is 88.14%, which is higher than the first health document. After that this research confirm the results with ROUGE. ROUGE compare the original health ethics protocol document with the data that has been extracted with RegEx, where the results can be seen in Table I, where P is Precision, R is Recall, and F is F-score. From the ROUGE result, it is confirmed that the RegEx extraction result from the second health ethics protocol document is more similar with the original document than the RegEx extraction result from the first health ethics protocol document.

After all the content of the ethical protocol has been extracted with RegEx, the next step is pre-processing it to remove noise in the content. Several techniques were carried out, namely (1) Case Folding, (2) Tokenization, (3) Stopwords Removal, and (4) Stemming. As in Table II, a sentence is processed to get the important words without any conjunctions, punctuation, spaces, etc. The extraction and pre-process stages in this study are very important, because they can determine whether the results of a summary are correct or deviate.

After the tokenization and stopword removal processes are carried out, we can understand what the important content written by the respondent refers to by calculating the intensity of each word. In Fig 2 and Fig 3 it can be seen that respondents in the health research ethics protocols document entitled "Studi Kasus Analisis Faktor Kepatuhan Karyawan" focused more on research related to compliance, implementation, vigilance, officers and hospitals.

After all content is pre-processed, cosine similarity is calculated for each token in the document. The results of this similarity calculation will be summarized using MMR and TextRank technique. The example summary extracted with MMR dan TextRank and the summary written by the respondents can be seen in Table III. Based on the extracted summary on Table III, we evaluate using the ROUGE Toolkit with the types ROUGE-2, ROUGE-3, ROUGE-4, ROUGE-L, and ROUGE-W which can be seen in Table IV.

The ROUGE evaluation results of extracted summary in Table IV will be normalized with the ROUGE results from the data extracted with RegEx with the original ethical protocol document as in Table I, because the extracted document with RegEx is not 100% the same as the original protocol document and also because the summary results come from extracted data with RegEx not from the original protocol document. Normalization of ROUGE results can be seen in Table V. Equation 3 is used to normalized the ROUGE value, where ROUGE of summary need to be divided by ROUGE of extracted data to get the value of normalized ROUGE.

$$ROUGE_{Norm} = \frac{(ROUGE_{sum} \times ROUGE_{extractedData})}{100} \quad (3)$$

TABLE III. THE EXAMPLE OF DOCUMENT SUMMARY AND COMPARISON WITH ORIGINAL SUMMARY

| Document | Summary |
|---|---|
| I | **Summary written by respondent** |
| | Tujuan peneliti ini adalah untuk menganalisis faktor organisasi, faktor individu dan faktor pekerjaan yang mempengaruhi kepatuhan petugas dalam penerapan prinsip kewaspadaan standar di rumah sakit. Penelitian ini menggunakan metode cross sectional, dengan pengambilan sampel secara simple random sampling sejumlah 89 orang petugas RSUD dr. R. Soedjono Selong. Pengumpulan data dilakukan dengan cara pengisian kuisioner dan observasi. sedangkan data pendukung dikumpulkan melalui studi dokumentasi dan wawancara mendalam. Analisis data dengan uji statistik SPSS. Alamat * Penelitian ini dilakukan karena terdapat kasus infeksi tuberkulosis dan hepatitis B pada petugas di RSUD dr. R. Soedjono dan hasil audit terhadap kepatuhan kewaspadaan standar yang rendah. manfaatnya adalah hasil penelitian ini dapat menjadi masukan bagi instansi mengenai pelaksanaan program di rumah sakit, menjadi bahan masukan untuk perencanaan program upaya pencegahan dan pengendalian infeksi bagi petugas, serta menjadi bahan untuk intervensi upaya pengendalian kejadian kecelakaan dan penyakit akibat kerja di rumah sakit. |

| | |
|---|---|
| | **Summary extracted using TextRank** |
| | Dekat dengan ibukota propinsi jenis penelitian ini adalah penelitian dengan pendekatan penelitian kuantitaf. Analisis regresi yang digunakan dalam penelitian ini adalah ujiberdasarkan aspek pengumpulan data, penelitian ini adalah penelitian observasional karena pada penelitian ini hanya dilakukan pengamatan, tanpa memberikan perlakuan terhadap obyek penelitian. Kriteria exlusi adalah jajaran manajemen, administrasi perkantoran dan petugas sim rs penelitian terhadap responden dihentikan apabila resonden tidak setuju untuk mengisi petugas rumah sakit yang terlibat sebagai calon subyek atau responden penelitian akan mendapatkan penjelasan tentang pelaksanaan penelitian. 2. Rancang bangun penelitian rancang bangun penelitian dengan menggunakan pendekatan cross sectional, yaitu penelitian yang desain pengumpulan datanya dilakukan pada satu titik (at one point in time); Fenomena yang diteliti adalah selama satu periode pengumpulan data. Lokasi penelitian adalah rumah sakit milik pemerintah daerah kab. Selanjutnya, responden penelitian dimohon untuk mengisi dan menandatangani lembar persetujuan setelah penjelasan (informed consent), yaitu dengan menyatakan bahwa bersedia atau tidak bersedia untuk secara sukarela menjadi subyek/responden penelitian dengan penuh kesadaran serta tanpa keterpaksaan. Pengambilan data penelitian melalui lembar kuesioner dan observasi langsung, yang dilakukan secara bertahap, dengan waktu pelaksanaan yang terlebih dahulu diagendakan bersama responden, agar tidak mengganggu aktivitas tidak ada responden dibawah umur cinderamata bagi responden penelitian, yaitu bolpoin dan paket alat kebersihan diri (sabun mandi, sikat gigi, pasta gigi, shampo dan handuk kecil) akan dilakukan presentasi hasil penelitian di rsud dr. Soedjono dengan mengundang pihak manajemen dan perwakilan petugas informasi kerahasiaan data responden dengan menggunakan kode nomer kode nomer dalam pengolahan dan analisis data. Analisi data yang digunakan dalam penelitian tersebut adalah : a. Analisis univariat terhadap tiap variabel dari hasil penelitian dengan menggunakan tabel distribusi frekuensi sehingga menghasilkan distribusi dan presentase dari setiap b. Analisis bivariat ditujukan untuk mengetahui ada tidaknya hubungan antar variabel bebas yaitu variabel lack of control dengan variabel basic causes, serta mengetahui hubungan antara variabel bebas dengan variabel terikat, yaitu variabel basic causes dengan variabel immediate causes. Analisis regresi dalam penelitian ini adalah analisis pengaruh faktor organisasi, faktor pekerjaan pengetahuan, sikap, motivasi, dan kepatuhan penerapan kewaspadaan standar. |
| | **Summary extracted using MMR** |
| | Analisis regresi yang digunakan dalam penelitian ini adalah uji. Kriteria exlusi adalah jajaran manajemen, administrasi perkantoran dan petugas sim rs penelitian terhadap responden dihentikan apabila resonden tidak setuju untuk mengisi petugas rumah sakit yang terlibat sebagai calon subyek atau responden penelitian akan mendapatkan penjelasan tentang pelaksanaan penelitian. Rumah sakit ini terletak dikota selong dan merupakan rumah sakit kelas c. Rsud dr. r. Soedjono selong saat ini ter-akreditasi madya rumah sakit sudah memiliki pengamanan evakuasi kondisi bencana, pemadam kebakaran dan kondisi darurat lainnya lokasi berada di tengah kota selong. Pengolahan data hasil penelitian dilakukan dengan menggunakan program spss for windows melalui beberapa tahapan yaitu editing, coding, entry data dan pengolahan data. Analisi data yang digunakan dalam penelitian tersebut adalah : a. Analisis univariat terhadap tiap variabel dari hasil penelitian dengan menggunakan tabel distribusi frekuensi sehingga menghasilkan distribusi dan presentase dari setiap b. Analisis bivariat ditujukan untuk mengetahui ada tidaknya hubungan antar variabel bebas yaitu variabel lack of control dengan variabel basic causes, serta mengetahui hubungan antara variabel bebas dengan variabel terikat, yaitu variabel basic causes dengan variabel immediate causes. Berdasarkan aspek pengumpulan data, penelitian ini adalah penelitian observasional karena pada penelitian ini hanya dilakukan pengamatan, tanpa memberikan perlakuan terhadap obyek penelitian. Analisis regresi dalam penelitian ini adalah analisis pengaruh faktor organisasi, faktor pekerjaan pengetahuan, sikap, motivasi, dan kepatuhan penerapan kewaspadaan standar. Pengambilan data penelitian melalui lembar kuesioner dan observasi langsung, yang dilakukan secara bertahap, dengan waktu pelaksanaan yang terlebih dahulu diagendakan bersama responden, agar tidak mengganggu aktivitas tidak ada responden dibawah umur cinderamata bagi responden penelitian, yaitu bolpoin dan paket alat kebersihan diri (sabun mandi, sikat gigi, pasta gigi, shampo dan handuk kecil) akan dilakukan presentasi hasil penelitian di rsud dr. Soedjono dengan mengundang pihak manajemen dan perwakilan petugas informasi kerahasiaan data responden dengan menggunakan kode nomer kode nomer dalam pengolahan dan analisis data. |

TABLE IV. EVALUATION OF EXTRACTED SUMMARY WITH ROUGE

| Data | ROUGE Score | | | | |
|---|---|---|---|---|---|
| | 2 | 3 | 4 | L | W |
| Doc 1 (Text Rank) | P: 4.35<br>R: 9.09<br>F: 5.88 | P: 0.34<br>R: 0.70<br>F: 0.45 | P: 0.00<br>R: 0.00<br>F: 0.00 | P: 10.00<br>R: 20.83<br>F: 13.51 | P: 10.00<br>R: 20.83<br>F: 13.51 |
| Doc 2 (Text Rank) | P: 17.22<br>R: 19.19<br>F: 18.15 | P: 8.64<br>R: 9.63<br>F: 9.11 | P: 4.67<br>R: 5.20<br>F:4.92 | P: 21.45<br>R: 23.90<br>F: 22.61 | P: 21.45<br>R: 23.90<br>F: 22.61 |
| Doc 1 (MMR) | P: 7.36<br>R: 52.80<br>F: 33.78 | P: 1.68<br>R: 3.52<br>F: 2.27 | P: 0.67<br>R: 1.42<br>F: 0.91 | P: 11.67<br>R: 24.31<br>F: 15.77 | P: 11.67<br>R: 24.31<br>F: 15.77 |
| Doc 2 (MMR) | P: 17.79<br>R: 19.56<br>F: 18.63 | P: 10.44<br>R: 11.48<br>F: 10.93 | P: 6.76<br>R: 7.43<br>F: 7.08 | P: 21.07<br>R: 23.16<br>F: 22.07 | P: 21.07<br>R: 23.16<br>F: 22.07 |

TABLE V. NORMALIZED ROUGE OF EXTRACTED SUMMARY

| Data | ROUGE Score | | | | |
|---|---|---|---|---|---|
| | 2 | 3 | 4 | L | W |
| Doc 1 (Text Rank) | P: 1.19 | P: 0.09 | P: 0.00 | P: 3.05 | P: 3.05 |
| | R: 2.45 | R: 0.18 | R: 0.00 | R: 6.30 | R: 6.30 |
| | F: 1.60 | F: 0.11 | F: 0.00 | F: 4.10 | F: 4.10 |
| Doc 2 (Text Rank) | P: 4.70 | P: 2.22 | P: 1.16 | P: 6.55 | P: 6.55 |
| | R: 5.18 | R: 2.45 | R: 1.27 | R: 7.22 | R: 7.22 |
| | F: 4.93 | F: 2.33 | F: 1.21 | F: 6.87 | F: 6.87 |
| Doc 1 (MMR) | **P: 4.35** | **P: 0.91** | **P: 0.34** | **P: 6.32** | **P: 6.32** |
| | **R: 31.08** | **R: 1.90** | **R: 0.71** | **R: 13.13** | **R: 13.13** |
| | **F: 19.92** | **F: 1.23** | **F: 0.46** | **F: 8.53** | **F: 8.53** |
| Doc 2 (MMR) | **P: 10.51** | **P: 5.65** | **P: 3.40** | **P: 11.42** | **P: 11.42** |
| | **R: 11.51** | **R 6.20** | **R: 3.73** | **R: 12.51** | **R: 12.51** |
| | **F: 10.98** | **F 5.92** | **F: 3.56** | **F: 11.91** | **F: 11.91** |

This result showed that the extracted summary is depends on the quality of the extracted results with RegEx which is input from ATS. In the second protocol document, the

extraction results with RegEx have a similarity of 88.14% with the original protocol document, so that the summary results with TextRank and MMR provide a good ROUGE value, while in the first protocol document, the extraction results with RegEx have a similarity of 86.51% with the original protocol document, so that the summary results with TextRank and MMR give a worse ROUGE value than the ROUGE of the second document summary. From Table V it can also be seen that the low ROUGE value can also be caused by linguistic factors in the language, in the second document, the grammar structure of the language is better than the first document, so the extracted summary is also better. All extracted summary, both with Text Rank and MMR have ROUGE-2 value greater than ROUGE-3 and ROUGE-4, this proves that Text Rank and MMR can choose good word pairs or overlapping bigrams. Compare to TextRank, ROUGE-2 of MMR is greater, this also proves that MMR is better than TextRank in choosing word pairs or overlapping bigrams.

The comparison of summary results with TextRank and MMR shows that MMR produces better summary results than TextRank. Summary of the first and second protocol document generated by MMR produced higher F-score, Precision and Recall from ROUGE-2, ROUGE-3, ROUGE-4, ROUGE-L, and ROUGE-W than Text Rank.

As mention before, when compared to TextRank, the performance of MMR is better, but the F-score of MMR are still less than 50 because there are still many irrelevant words chosen by MMR in the generated summary and the summary generated by MMR are also influenced by the quality of the text in the original document. Therefore, in the future, the summarization will be carried out using abstractive methods like deep learning, which will generate new sentences that are more relevant from the document.

## V. CONCLUSION

In this research, we have built a system to perform summary of research ethics protocol automatically using two summarization methods, namely Text Rank and MMR. The MMR method is considered to have better results than TextRank based on the ROUGE evaluation results. Then the content extraction process is a very important process in this research. With the extraction result of 86.51% in document 1, it produces an F-score value of 19.92 using MMR. While the extraction of document 2 of 88.14% produces an F-score of 10.98 using MMR. For future work, this research will use abstractive methods to generate a better summary. For future work we will make the dataset larger by collecting more health research ethics documents and possibly dividing them into smaller sub-documents, after that we will use the larger dataset and train them with sequence to sequence (seq2seq) model to achieve better results by finding relevant features in the written text.

## REFERENCES


[1] World Health Organization, "Working together for health: The World Health Report 2006. Policy brief," Geneva, 2006.

[2] World Health Organization, *International Ethical Guidelines for Health-related Research Involving Humans Prepared by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO)*. 2016.

[3] W. M. Association, "World Medical Association declaration of Helsinki: Ethical principles for medical research involving human subjects," *JAMA - Journal of the American Medical Association*, vol. 310, no. 20. American Medical Association, hal. 2191–2194, Nov 2013, doi: 10.1001/jama.2013.281053.

[4] KEMENTERIAN KESEHATAN REPUBLIK INDONESIA, "Pedoman dan Standar Etik Penelitian dan Pengembangan Kesehatan Nasional," *Kementeri. Kesehat. RI*, hal. 1–158, 2017.

[5] W. S. El-Kassas, C. R. Salama, A. A. Rafea, dan H. K. Mohamed, "Automatic text summarization: A comprehensive survey," *Expert Syst. Appl.*, vol. 165, Mar 2021, doi: 10.1016/j.eswa.2020.113679.

[6] H. Oliveira *et al.*, "Assessing shallow sentence scoring techniques and combinations for single and multi-document summarization," *Expert Syst. Appl.*, vol. 65, hal. 68–86, Des 2016, doi: 10.1016/j.eswa.2016.08.030.

[7] P. M. Sabuna dan D. B. Setyohadi, "Summarizing Indonesian text automatically by using sentence scoring and decision tree," *Proc. - 2017 2nd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2017*, vol. 2018-Janua, hal. 1–6, Feb 2018, doi: 10.1109/ICITISEE.2017.8285473.

[8] D. Gunawan, S. H. Harahap, dan R. Fadillah Rahmat, "Multi-document Summarization by using TextRank and Maximal Marginal Relevance for Text in Bahasa Indonesia," *Proceeding - 2019 Int. Conf. ICT Smart Soc. Innov. Transform. Towar. Smart Reg. ICISS 2019*, Nov 2019, doi: 10.1109/ICISS48059.2019.8969785.

[9] I. M. S. Putra, Y. Adiwinata, D. P. S. Putri, dan N. P. Sutramiani, "Extractive Text Summarization of Student Essay Assignment Using Sentence Weight Features and Fuzzy C-Means," *Int. J. Artif. Intell. Res.*, vol. 5, no. 1, hal. 13–24, Jun 2021, doi: 10.29099/IJAIR.V5I1.187.

[10] D. Gunawan, C. A. Sembiring, dan M. A. Budiman, "The Implementation of Cosine Similarity to Calculate Text Relevance between Two Documents," *J. Phys. Conf. Ser.*, vol. 978, no. 1, hal. 012120, Mar 2018, doi: 10.1088/1742-6596/978/1/012120.

[11] I. G. M. Darmawiguna, G. A. Pradnyana, dan I. B. Jyotisananda, "Indonesian sentiment summarization for lecturer learning evaluation by using textrank algorithm," *J. Phys*, hal. 12024, 2021, doi: 10.1088/1742-6596/1810/1/012024.

[12] H. N. Fejer dan N. Omar, "Automatic multi-document Arabic text summarization using clustering and keyphrase extraction," *J. Artif. Intell.*, vol. 8, no. 1, hal. 1–9, 2015, doi: 10.3923/jai.2015.1.9.

[13] C.-Y. Lin, "ROUGE: A Package for Automatic Evaluation of Summaries."